MMSPICE: A PHYSICAL ALTERNATIVE TO GUMMEL-POON/SPICE AND A SEMI-NUMERICAL MIXED-MODE SIMULATOR FOR ADVANCED BIPOLAR TECHNOLOGY CAD

By

HANGGEUN JEONG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1989

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to Professor Jerry Fossum for having given me an opportunity to work as one of his privileged graduate students on interesting research topics. His devoted and methodical guidance in technical matters, support in other nontechnical areas, and continuous encouragement were crucial through all phases of this work. I also would like to thank the other members of my supervisory committee, Professors Lindholm, Burk, Eisenstadt, and Kurzweg, for their willingness to serve on my committee.

I am also thankful to numerous people I have interacted with during my years of graduate study. First I am grateful to Mr. Dan FitzPatrick for implementing the model in the SPICE2 program. Thanks are also extended to many of my friends who helped me through technical discussions or by cheering me up in difficult times. I cannot mention all of them, but I should mention Drs. Surya Veeraraghavan, Young-June Yu and Messrs. Myung-Suk Jo, Yeong-Seuk Kim, Jin-Young Choi, Joohyun Jin. My deepest gratitude goes to my parents and parents-in-law. Last but not the least I thank my wife Hyo for her support through long years of my graduate study. I also acknowledge the financial support of the Semiconductor Research Corporation.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS ii

ABSTRACT v

CHAPTERS

1 INTRODUCTION 1

2 PHYSICAL MODELING OF COLLECTOR HIGH-CURRENT EFFECTS .. 10

2.1 Introduction 10
2.2 Qualitative Analysis 11
2.3 Physical Model 19

3 SEMI-NUMERICAL CHARGE-BASED BIPOLAR TRANSISTOR MODEL. 28

3.1 Introduction 28
3.2 Model Development 29
3.2.1 Base Analysis 31
3.2.2 Collector Analysis 34
3.2.3 Boundary Conditions and Linkages 40
3.2.4 Charging Currents 46
3.2.5 Model Network Representation 49
3.3 Experimental Support 54
3.3.1 Parameter Extraction 55
3.3.2 Simulations 62

4 PHYSICAL MODELING OF WEAK IMPACT IONIZATION 69

4.1 Introduction 69
4.2 Physical Model 71
4.3 Experimental Support 76

5 MMSPICE DEVELOPMENT 87

5.1 Introduction 87
5.2 Model Integration 90
5.3 Numerical Technique 94
5.4 SPICE Code Modification 96
5.5 Demonstration/Benchmarking 100

6 SUMMARY AND SUGGESTIONS FOR FUTURE WORK 112

APPENDICES

A DETAILED MODEL ALGORITHM 114

B PARAMETER OPTIMIZATION/EXTRACTION 123

C MMSPICE BENCHMARK CIRCUITS 132

REFERENCES 138

BIOGRAPHICAL SKETCH 142

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

MMSPICE: A PHYSICAL ALTERNATIVE TO GUMMEL-POON/SPICE AND A SEMI-NUMERICAL MIXED-MODE SIMULATOR FOR ADVANCED BIPOLAR TECHNOLOGY CAD

By

HANGGEUN JEONG

December 1989

Chairman: Dr. J. G. Fossum
Major Department: Electrical Engineering

This dissertation is concerned with physical, semi-numerical modeling of advanced bipolar transistors (BJTs) for device and circuit simulation, and the development of a semi-numerical mixed-mode simulator for bipolar and BiCMOS technologies. A new charge-based model for the advanced polysilicon self-aligned BJT has been developed based on one-dimensional regional analyses of the carrier transport that account for high-current effects, impact ionization, and nonreciprocal capacitances. The base regional analysis yields, for all injection levels, the transport current in terms of quasi-neutral-base boundary (QNB) conditions. The collector regional analysis, which accounts for both ohmic and nonohmic quasi-saturation, defines all possible charge conditions that determine the physical linkages to the base

transport via proper moving QNB boundary conditions. The linked regional analyses, which require numerical iterative solutions, characterize transport currents and quasi-static stored charges (and charging currents) in terms of device technological parameters that relate directly to the device structure. Thus, a semi-numerical mixed-mode device/circuit simulation is enabled. Weak impact ionization in the high-field epitaxial-collector region is numerically evaluated based on the electric field distribution provided by the collector transport analysis. The resulting model is implemented in the circuit simulator SPICE2 to create MMSPICE, a physical alternative to Gummel-Poon/SPICE and a semi-numerical mixed-mode simulator. Unique implementation problems, e.g., techniques for numerically solving model equations and for ensuring current/charge continuity, are addressed. Simulated dc characteristics for typical advanced BJTs are compared against corresponding measurements to provide some experimental support for the model. Utility of MMSPICE in technology-CAD applications is demonstrated by ECL transient simulations. MMSPICE is benchmarked against SPICE for a variety of bipolar circuits. Algorithms for parameter optimization are described and their use in parameter extraction is illustrated.

CHAPTER 1
INTRODUCTION

Advances in silicon bipolar technology in the last decade have led to significant improvements in bipolar transistor (BJT) performances, especially in switching speed. These improvements were obtained mainly by scaling the device both vertically and horizontally. This scaling was achieved through a self-aligned device structure with polysilicon contacts, typically fabricated by double-polysilicon processes [Zde87] (see Fig. 1.1). Use of a polysilicon contact for the emitter made possible vertical scaling without degrading the transistor current gain. Use of a polysilicon base contact was instrumental in effecting drastic reductions in the parasitics associated with the extrinsic base-collector junction and in the extrinsic base resistance by enabling the self-alignment of the emitter to the base, as well as the placement of the base electrode over the field oxide.

Despite the impressive progress made in the bipolar technology, computer simulation tools, which are essential to the optimization of device and circuit designs for the technology, have not kept pace with it. In integrated circuit

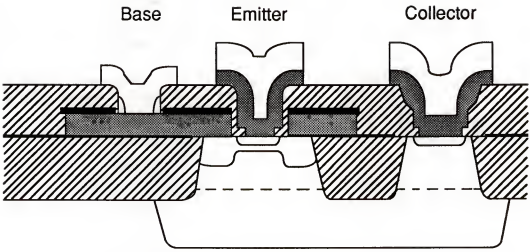


Fig. 1.1 Cross-sectional view of a typical advanced bipolar transistor fabricated by double-polysilicon process. Typical width of the emitter is less than 1 μm.

(IC) development and manufacturing today, a technology-CAD (TCAD) system, which comprises process, device, and circuit simulators, is essential for exploring alternative designs and evaluating various tradeoffs without time-consuming and costly fabrications. An effective TCAD system requires not only reliable constituent simulation tools, but also proper integration of such tools. For the latter requirement, there is growing interest in a mixed-mode device/circuit simulator which can be used in TCAD work to predict the circuit performances based on the specified device structure. However, existing mixed-mode simulators [Eng82, May87, Rol88], which are purely numerical, are too CPU-intensive to be used for statistical simulations, e.g., for design for manufacturability. A more computationally efficient alternative is now needed to complement the conventional mixed-mode simulators.

For circuit simulation, improved BJT device models, implemented in existing circuit simulators, are needed to keep pace with the advancing bipolar technology. The Gummel-Poon [Get78] BJT model, which has been the workhorse for bipolar circuit simulation, is too empirical to properly simulate advanced-structure BJTs. The model was derived based on the integral charge-control relation (ICCR) which relates the transport current to the integral of majority-carrier

density in the base region. This ICCR itself is physically valid in most regions of transistor operation, except hard saturation. However, no explicit expression can analytically be derived for the majority-carrier integral. In the Gummel-Poon model, empirical parameters are introduced to enable an explicit representation of the integral. The explicit formalism facilitates a numerically efficient implementation in a circuit simulator, while the empirical (non-technological) nature of the model parameters renders the model usable even when no information is available with regard to the device structure being simulated. However, these advantages are achieved at the expense of losing some of the physical basis in the ICCR. For example, the Gummel-Poon model accounts only empirically for high-current effects which have become more pronounced in the scaled advanced BJTs.

To remove these deficiencies of the Gummel-Poon model, and to provide a computationally efficient alternative to conventional mixed-mode simulations, a new physics-based BJT model should be developed, assuming the availability of device structural information for the extraction of model parameters. This assumption would not seriously restrict the model utility because such technological information is available in typical IC-development environments. In deriving

this model, carrier transport equations can be greatly simplified by restricting model applicability to a specific type of structure and by resorting to a regional modeling approach. Unlike the conventional compact device models in circuit simulators, this application-specific model can be formulated using an implicit formalism to semi-numerically, but physically account for interactions between the regional transports. Inclusion of technological parameters enables not only more accurate accountings for device physics, but also a predictive modeling whereby the effects of variation in technological parameters on circuit performances can be simulated. Hence, such a semi-numerical BJT model can serve the dual purpose of improving the accuracy of the device model for circuit simulation and of enabling an application-specific, but less CPU-intensive mixed-mode device/circuit simulation for bipolar and BiCMOS TCAD.

To develop such a physics-based BJT model, regional analyses must first be formulated for the generic advanced BJT structure, and then their solutions must be physically linked. Recently, novel approaches for physically modeling the base and collector carrier transports were reported. Kull et al. [Kul85] developed a compact physical model for quasi-saturation effects. De Graaff and Kloosterman [Gra85] demonstrated a new formulation of the base transport that can

physically account for the effects of built-in field and high-injection. However Kull's model is derived based on the assumption that the entire epitaxial collector is quasi-neutral. This assumption makes the model empirical when any space-charge region (SCR) exists in the epitaxial collector. This empiricism also makes it difficult to establish physical linkages between the base and collector transports. In De Graaff's model, the boundary conditions are treated empirically, partially negating the advantages of using a physical modeling approach for the base transport.

By extending and refining these regional analyses and linking them based on physical interactions between them, a new physical BJT model is derived. The new model is implicitly defined by a set of nonlinear equations, and hence a numerical solution is required. Solution of the model equations yields expressions for the stored regional charges, as well as for the transport current. Some of the regional charges are shared between the terminals and need to be partitioned between the sharing terminals. Derivatives of these regional charges with proper partitioning represent charging currents in the charge-based BJT model. Implementing this kind of model in a circuit simulator poses unique problems which are not encountered in implementing conventional models. These problems arise from the implicit

nature of the model and complex modes of physical linkages between regional analyses. Such problems must be adequately addressed for this new modeling approach to be viable.

This dissertation addresses these problems; it is concerned with the development and implementation of a physics-based model for the advanced bipolar transistor, which can be used not only for an accurate circuit simulation, but also for a computationally efficient (semi-numerical) mixed-mode device/circuit simulation. The major contributions made in this work are as follows:

(1) formulation of a comprehensive collector transport based on physical analyses of high-current effects (quasi-saturation);

(2) development of a charge-based semi-numerical BJT model based on physical linkages between base and collector transports;

(3) modeling of weak impact ionization based on bias-dependent electric field distribution;

(4) implementation of the developed model in the source code of the circuit simulator SPICE2 to create MMSPICE-1.1, including benchmarking and demonstration of the new semi-numerical mixed-mode device/circuit simulator.

In Chapter 2 a new physical analysis of collector high-current effects is described. This analysis extends the

previous work [Kul85] by accounting for the possible formation of a current-induced space-charge region in the epitaxial collector (away from the junction) at very high currents. Thus, the new analysis physically accounts for nonohmic quasi-saturation, as well as ohmic quasi-saturation.

In Chapter 3, the base transport is modeled by solving an ambipolar transport in the quasi-neutral base. Then the quasi-saturation model derived in Chapter 2 is extended to define a comprehensive collector model. The resulting base and collector transports are linked based on the physical nature of their interactions as determined by the operating condition in the epitaxial collector. The physically linked regional analyses yield the implicitly defined nonlinear model equations. The transient model is derived quasi-statically, using the charge-based formalism with some of the charges partitioned to account for the non-quasi-static delay in building up or discharging those charges. To provide experimental support, dc simulations and measurements are presented. Parameter extraction methods used for the simulations are also discussed.

In Chapter 4 a new physical way of accounting for weak impact ionization is described. The new approach is based on the numerical integration of the field-dependent ionization coefficient in the high-field region in the collector. The

field distribution is provided by the semi-numerical BJT model described in Chapter 3. The model is modified to include the impact-ionization current. Simulations are done for dc characteristics and the results are discussed with regard to the experimental data.

In Chapter 5 the implementation of the developed charge-based BJT model into SPICE2 is discussed, with emphasis on the implementation problems associated with the unique nature of the model. Representative simulations and benchmarks are presented.

In Chapter 6 the main accomplishments of this dissertation are summarized and areas for future work are suggested.

Appendix A details the model algorithms as used in the MMSPICE-1.1 implementation.

Appendix B describes the model parameter optimization/extraction algorithm and its implementation in Unix-system environments.

Appendix C lists MMSPICE input files for the circuits used to benchmark MMSPICE-1.1.

CHAPTER 2
PHYSICAL MODELING OF COLLECTOR HIGH-CURRENT EFFECTS

2.1 Introduction

Today's scaled-down fast bipolar transistors tend to operate, for optimal circuit performances (e.g., minimum gate delay in ECL circuits), at current densities near or even exceeding the onsets of so-called high-current phenomena [Tas84]. The high-current effects in these submicrometer devices occur predominantly in the lightly doped epitaxial collector region. These effects are only empirically, and hence inadequately, accounted for in the existing bipolar transistor models for circuit simulators, e.g., the Gummel-Poon model in SPICE2 [Get78]. An improved model based on physical principles is now essential for accurate simulation of the high-current effects, whether they are incurred deliberately for optimal performances or inadvertently in improperly designed circuits.

Recently, Kull et al. [Kul85] extended the Gummel-Poon model to include quasi-saturation, or base push-out. Kull's compact extension, however, is not generally applicable because of the assumption that the entire epitaxial collector region is quasi-neutral. The model is useful for high-voltage

(thick-epi) transistors, but yields erroneous predictions for VLSI transistors when the current density is so high that a space-charge region must form in the collector region to support the current [Rey75].

As the first step in developing a physical BJT model for device/circuit simulation, Kull's model is extended in this chapter to account for the possible existence of the current-induced space-charge region in the epitaxial collector. This extension is facilitated by a qualitative analysis of the high-current phenomena, which is described in Section 2.2. In Section 2.3 the carrier transport problem in the induced space-charge region is solved analytically based on simplifying assumptions verified by numerical (BAMBI [Fra85]) simulations of typical transistors. Then, Kull's solution is modified to account for the moving boundary between the quasi-neutral region and the induced space-charge region.

2.2 Qualitative Analysis

The one-dimensional n^+-p-n-n^+ structure of an epitaxial bipolar transistor and typical I_C versus V_{CE} characteristics are shown in Fig. 2.1 and Fig. 2.2, respectively. Two-dimensional flow associated with high current densities in the epitaxial collector is neglected [Kul75]. Current flow in the epitaxial collector region produces two effects: a

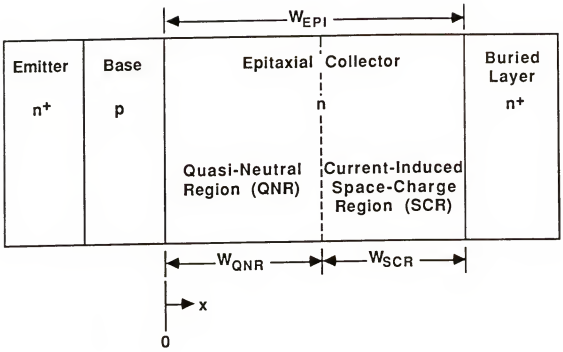


Fig. 2.1 One-dimensional structure of an n^+-p-n-n^+ epitaxial bipolar transistor, including indication of a current-induced space-charge region in the epitaxial collector.

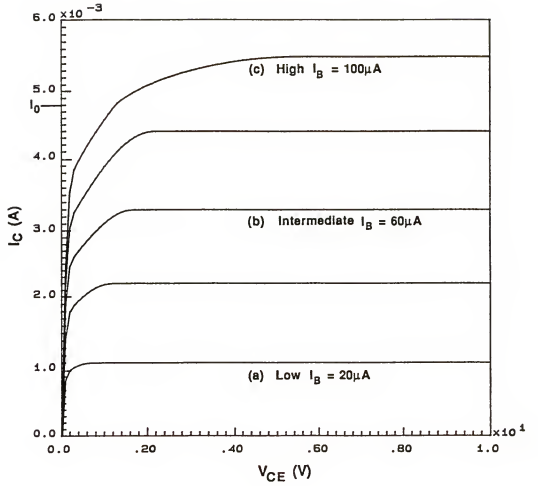


Fig. 2.2 I_C versus V_{CE} characteristics for a typical n^+-p-n-n^+ bipolar transistor, with the value of I_O indicated.

modulation of charge in the collector-base junction space-charge region, and an ohmic drop in the quasi-neutral collector region. Previous studies [Pal69],[Whi69],[Bow73] have indicated under what conditions one effect is dominant over the other in producing the ultimate forward biasing of the collector-base junction at sufficiently high current densities. "Base push-out" is usually attributed to the space-charge modulation effect, and "quasi-saturation" to the ohmic drop.

Rey et al. [Rey75] presented a unified treatment of the two seemingly distinct mechanisms by focusing on the electric field profile in the "collecting zone." The key hypothesis was that the collecting zone should become space-charged when the majority-carrier drift velocity saturates. Based on this unified view of the high-current phenomena, the following regions of operation are defined.

a) Ohmic Quasi-Saturation (OQS): The electric field in the collecting zone is low enough that the carrier mobility is independent of the electric field.

b) Pseudo-Ohmic Quasi-Saturation (POQS): The electric field in the collecting zone is high enough to cause some degradation of carrier mobility but not high enough to saturate the carrier velocity.

c) Non-Ohmic Quasi-Saturation (NOQS): The carrier drift velocity is effectively saturated in the collecting zone such that a space-charge region must form [Rey75].

Previous workers have used ill-defined terminology to describe OQS, POQS,NOQS, tending to ignore the different regions of operation and referring to "quasi-saturation" as "base push-out" or as the "Kirk effects". Consistent with the physical nature of these effects and their connections [Rey75], the term "quasi-saturation" is used here to encompass all the high-current effects that occur in the collector region and degrade current gain and frequency response, but OQS is distinguished from NOQS.

The electric field required for velocity saturation is not unambiguously defined because the saturation occurs gradually as the field increases. Numerical simulations (Section 2.3) show that the space-charge region begins to form with the electric field well below the value commonly accepted for saturation (~ 50 kV/cm) of the electron velocity. This suggests that the simplified piecewise-linear velocity field model illustrated in Fig. 2.3 is suitable for a transistor model intended for circuit simulations. The electron mobility is assumed to be constant (μ_{no}) for electric fields below E_S, above which the electron velocity is constant (saturated at v_S). The implicit neglect of mobility

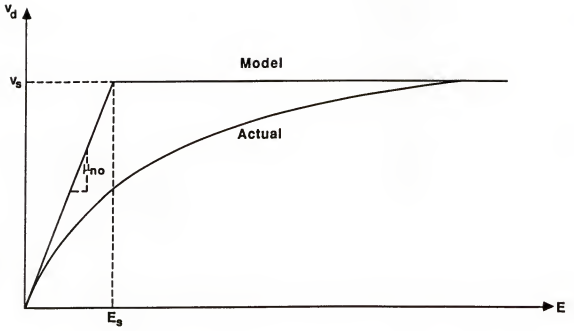


Fig. 2.3 Simple model for carrier drift velocity-electric field dependence, referenced to the actual dependence. Note that $v_d = \mu_{no}E$ for $E < E_S$.

degradation in the POQS region results in some uncertainty regarding the boundary between quasi-saturation and the forward-active mode, but the error in terms of I_C is not significant for most applications.

Based on this simplified mobility model, an insightful qualitative description of the I_C versus V_{CE} characteristics can be given. It is assumed that the current gain is high enough to neglect minority carrier current in the epitaxial collector region. Three representative $I_C(V_{CE})$ curves, corresponding to different base drives, as shown in Fig. 2.2, illustrate the description.

a) Low I_B (or V_{BE}): the collector current in the forward active mode, $I_{C(fa)}$, is much smaller than V_{BE}/R_{EPI}, where R_{EPI} is the unmodulated resistance of the epitaxial collector region. For this case, only OQS obtains, but is insignificant.

b) Intermediate I_B (or V_{BE}): $I_{C(fa)}$ is comparable to or much larger than V_{BE}/R_{EPI}, but smaller than I_O, where $I_O \equiv qAN_{EPI}v_S$ is the saturated-velocity electron drift current in the epitaxial collector. In this case, OQS is significant, and separates the forward-active mode from the saturation region as shown in Fig. 2.2. Since V_{CE} is approximately equal to the product of the electric field in the collecting zone and the collecting-zone width, an increase of V_{CE} must result in a

higher electric field and/or a wider collecting zone [Rey75]. Increased electric field in the collecting zone necessarily involves higher current, requiring decreased internal forward biasing of the collector-base junction (as characterized by the base transport analysis) and diminished conductivity modulation in the collector. Consequently, increases of electric field and of collecting-zone width occur simultaneously when V_{CE} increases. Finally, when V_{CB} (= V_{CE} - V_{BE}) exceeds $I_{C(fa)}R_{EPI}$, the collecting zone occupies the whole epitaxial layer and the device enters the forward-active mode.

c) High I_B (or V_{BE}): $I_{C(fa)}$ is larger than I_O, and hence NOQS obtains. As V_{CE} is increased, the behavior of the device is similar to that in b) until I_C exceeds I_O. When I_C is larger than I_O, an increase of the electric field in the collecting zone cannot, because of the carrier velocity saturation, support a higher current as required by the expansion of the collecting zone. Therefore, excess majority electrons are injected into the collecting zone, thereby creating a space-charge region. This provides the connection between OQS and NOQS as recognized by Rey et al. [Rey75]. For sufficiently high V_{CE}, the entire epitaxial layer becomes space-charged, and the device enters the forward-active mode.

2.3 Physical Model

Kull et al. [Kul85] derived a closed-form expression for collector current in terms of the majority-carrier quasi-Fermi potentials at the boundaries of the epitaxial collector region, accounting for, in fact OQS and POQS, but not for NOQS. Quasi-neutrality was assumed throughout the epitaxial collector region, thereby preventing proper modeling of NOQS. Kull's model accounts for conductivity modulation and field dependence of mobility (below velocity saturation) in a compact and accurate manner. Furthermore, the solution in terms of the majority-carrier quasi-Fermi potential can easily be connected to the base transport problem, making the interface with the Gummel-Poon model straightforward. However, to make the model valid for the NOQS region, it must be extended to account for the current-induced space-charge region, and the moving boundary between it and the quasi-neutral region (QNR) in the epitaxial collector (see Fig. 2.1).

To account for the possible existence of the SCR, a regional modeling approach is used, as depicted in Fig. 2.1. The presence of the SCR and its internal boundary depend on the current and voltage conditions. Let us consider two steady-state cases defined by W_{QNR}, the width of QNR, relative to W_{EPI}, the width of the n-epitaxial layer.

A. $W_{QNR} = W_{EPI}$

This case corresponds to Kull's model [Kul85] and is applicable as long as I_C is smaller than I_O:

$$I_C = \frac{K_1(V_{BCO}) - K_1(V_{BC}) - \ln\left[\frac{1 + K_1(V_{BCO})}{1 + K_1(V_{BC})}\right] + \frac{V_{BCO}-V_{BC}}{V_T}}{R_{EPI}/V_T} \qquad (2.1)$$

where $R_{EPI} \equiv W_{EPI}/(q\mu_{no}N_{EPI}A)$, $K_1(V) \equiv [1 + \gamma \exp(V/V_T)]^{1/2}$ (with V equal to V_{BCO} or V_{BC} here), $\gamma \equiv (2n_i/N_{EPI})^2$, $V_{BCO} \equiv [E_{Fn}(0) - E_{Fp}]/q$, $V_{BC} \equiv [E_{Fn}(W_{EPI}) - E_{Fp}]/q$, and $V_T \equiv kT/q$. In writing (2.1), the simplified mobility description (Fig. 2.2) is incorporated into Kull's result, which makes the resulting device model more suitable for circuit analysis without sacrificing much accuracy. In this case any space-charge in the transition region between the conductivity-modulated portion of the collector and the collecting zone is neglected following [Kul85].

B. $0 < W_{QNR} < W_{EPI}$

In this case, part of the epitaxial layer becomes space-charged as I_C exceeds I_O. The corresponding energy-band diagram and electric-field profile are shown in Fig. 2.4. To solve the carrier transport problem for this case, the

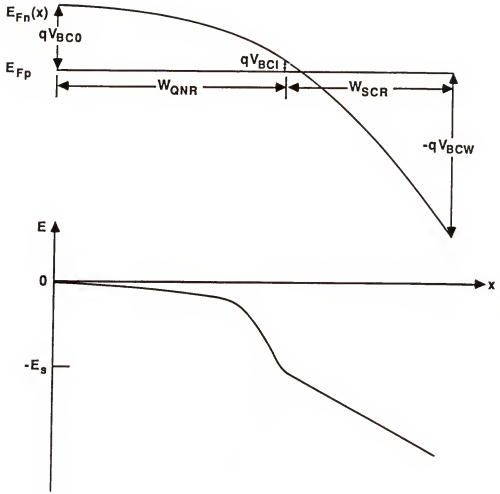


Fig. 2.4 Energy-band diagram and corresponding electric-field profile in the epitaxial collector region for the case $0 < W_{QNR} < W_{EPI}$. The transition region between the QNR and the SCR is assumed to be negligible.

following assumptions are made based on physical insight and numerical simulation results, which will be discussed later.

1) The induced SCR exists if and where $E > E_S$, and hence the majority-carrier velocity is saturated at v_S.

2) In the induced SCR, the majority-carrier current flows predominantly by drift. Then

$$I_C \approx qA(N_{EPI} + \Delta n)v_S \tag{2.2}$$

where Δn, the excess electron density in the SCR, is spatially constant.

3) The minority-carrier density in the SCR is negligible compared to the majority-carrier density ($N_{EPI} + \Delta n$).

Based on these assumptions, the carrier transport in the induced SCR is governed by the following form of Poisson's equation:

$$\frac{d^2E_{Fn}}{dx^2} = -\frac{q^2}{\varepsilon}\Delta n \tag{2.3}$$

where it is implicitly assumed that the electric field is defined by the gradient of the majority-electron quasi-Fermi level E_{Fn}. The boundary conditions associated with (2.3) are (see Fig. 2.4)

$$\frac{dE_{Fn}}{dx}\Big|_{x=W_{QNR}} = - qE_S \tag{2.4}$$

and

$$E_{Fn}(W_{EPI}) - E_{Fp} = qV_{BC}. \tag{2.5}$$

The differential equation (2.3) can be solved, using (2.2), (2.4), and (2.5), to yield the following expression for the SCR width:

$$W_{SCR} = -\frac{\varepsilon E_S}{q\Delta n} + [(\frac{\varepsilon E_S}{q\Delta n})^2 + \frac{2\varepsilon(\phi_C - V_{BC})}{q\Delta n}]^{1/2}. \tag{2.6}$$

where ϕ_C is the built-in potential for the base-collector junction. In deriving (2.6), the potential drop in the quasi-neutral region is neglected.

To describe the current-voltage characteristics for this case, the Kull's result (2.1) is applied to the shrunken QNR as implied by (2.6); that is

$$W_{QNR} = W_{EPI} - W_{SCR} \tag{2.7}$$

and V_{BCI}, which is the quasi-Fermi potential separation at $x = W_{QNR}$ (see Fig. 2.1), replace, respectively, W_{EPI} and V_{BC} in (2.1). First, however, V_{BCI} must be defined. Since any transition region between the SCR and the QNR is ignored, V_{BCI}

is defined to ensure continuity in the electron density at the boundary. Thus, at $x = W_{QNR}$, $n \approx N_{EPI} + \Delta n$, and because of the neutrality, $p \approx \Delta n$ at the boundary. Then

$$V_{BCI} = V_T \ln \frac{\frac{I_C}{qAv_S}\left(\frac{I_C}{qAv_S} - N_{EPI}\right)}{n_i^2} . \tag{2.8}$$

Thus, the modified version of (2.1) now yields

$$I_C = \frac{K_1(V_{BCO}) - K_1(V_{BCI}) - \ln\left[\frac{1 + K_1(V_{BCO})}{1 + K_1(V_{BCI})}\right] + \frac{V_{BCO}-V_{BCI}}{V_T}}{R'_{EPI}/V_T} \tag{2.9}$$

where $R'_{EPI} = W_{QNR}/qA\mu_{no}N_{EPI}$.

In the preceding analysis, it has been tacitly assumed that the transistor is not saturated, i.e., $V_{CE} > V_{BE}$. Although the general ambipolar transport analysis is complex, it can be noted that much of the saturation region can be treated by neglecting the minority-hole current in the collector region. Hence, (2.1) and (2.9) can be used to approximately characterize this region, except for hard saturation.

Quasi-static expressions for two components of stored excess charge in the collector region, which are needed to define the transient model, can be derived from the analysis.

In the SCR, $Q_{SCR} = -qAW_{SCR}\Delta n$, and in the QNR, Q_{QNR} is defined by the integral of $\Delta p(x)$ [Kul85] over the range $0 \leq x \leq W_{QNR}$:

$$Q_{SCR} = 0, \qquad \text{for } I_C \leq I_O$$

$$= -qAW_{SCR}\left(\frac{I_C}{qAv_S} - N_{EPI}\right), \qquad \text{for } I_C > I_O \tag{2.10}$$

and

$$Q_{QNR} = \frac{\frac{qA\gamma W_{QNR}}{4} N_{EPI}\left[\exp\left(\frac{V_{BCO}}{V_T}\right) - \exp\left(\frac{V_{BCI}}{V_T}\right)\right]}{I_C R'_{EPI}/V_T} - \frac{qAW_{QNR}n_i^2}{N_{EPI}} \tag{2.11}$$

Key assumptions used in our model extension have been checked using BAMBI, a two-dimensional device simulator. Fig. 2.5 shows selected results of BAMBI simulations of a virtual one-dimensional n^+-p-n-n^+ transistor structure in steady state at high I_C. The plots of the electron current density, the electron and hole densities, and the electric field reflect the SCR, characterized in accord with the model assumptions given prior to (2.3).

The preceding analysis of collector high-current effects does not constitute a complete collector transport model because the existence of the junction space-charge region is ignored. However, the quasi-saturation analysis developed

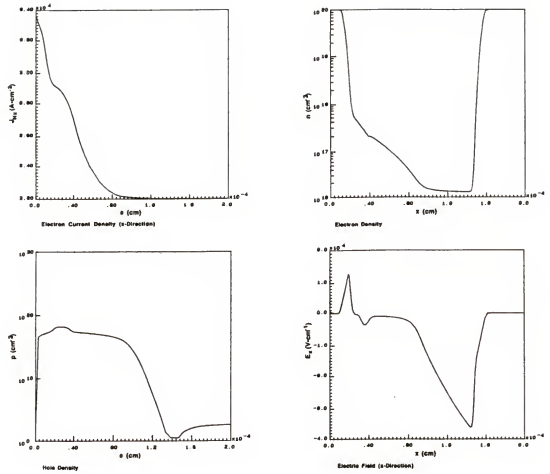


Fig. 2.5 Selected results of steady-state BAMBI simulations of a virtual one-dimensional n^+-p-n-n^+ transistor at high current. For these results, N_{EPI} = 1.0 X 10^{16} cm^{-3}, W_{EPI} = 1.0 μm; V_{CE} = 1.5V, V_{BE} =0.9 V. Note in the n(x) and E_x(x) plots the existence of the SCR (in the range 1.0μm < x < 1.4 μm), the characteristics of which support the model assumptions.

here serves as a basis in Chapter 3 for doing a comprehensive physical collector analysis as part of the development of a semi-numerical charge-based BJT model. Accordingly, Q_{SCR} defined in (2.10) will be generalized later to include the stored charge in the junction SCR, as well as that in the current-induced SCR.

CHAPTER 3
SEMI-NUMERICAL CHARGE-BASED BIPOLAR TRANSISTOR MODEL

3.1 Introduction

In Chapter 2, a physical model for the high-current effects in the collector was developed. In this chapter a physical model for the base transport is first derived. Then, the work in Chapter 2 is extended to derive a comprehensive collector model that facilitates physical linkages with the base transport analysis. The physically linked base-collector regional analyses define the semi-numerical physical charge-based BJT model.

The model avoids excessive empiricism, e.g., that implied by the transit times in previous models, and can be predictive; key parameters are technological. The details of the model developments and physical linkages between the base and collector transports are described in Section 3.2. Solutions of ambipolar transports coupled with moving boundary conditions in the base and collector regions yield not only the steady-state current-voltage characteristics, but also the expressions for the stored charges. The charge-based formalism used to derive the transient model is more representative of the (distributed) charge dynamics than the

conventional capacitance-based models such as the Gummel-Poon model. Time derivatives of the stored charges can account for the inherent nonreciprocal transcapacitances of the BJT. Certain charges shared by more than one terminal are partitioned between the sharing terminals to give physical reality to the model [Fos86].

In Section 3.3, the model is verified by dc measurements of contemporary test devices fabricated at Motorola, Inc. [Zde87]. Parameters are extracted from layout and process information as well as from measured data. Due to the physical nature of the model, parameter extractions are straightforward, and the extracted parameters are applicable over a wide range of operating conditions.

3.2 Model Development

In this development, an n^+-p-n-n^+ epitaxial bipolar transistor fabricated by contemporary double-polysilicon self-aligned processes [Zde87] is considered. The one-dimensional structure of the BJT is illustrated in Fig. 3.1 with its regional boundaries labeled and regional charges indicated. A regional modeling approach is used to analytically characterize the device transport. First, the device is partitioned into several regions by identifying where an approximate charge-neutrality prevails (QNR) and

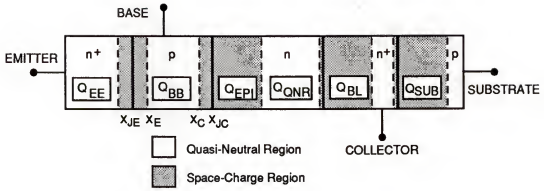


Fig. 3.1 One-dimensional structure of an n^+-p-n-n^+ epitaxial bipolar transistor with its regional boundaries labeled and regional charges indicated.

where it does not (SCR). Then, ambipolar transport problems are solved in the QNRs through simplifying approximations to obtain analytic expressions for the transport currents and stored charges in terms of the moving boundary locations and carrier densities at the boundaries. These boundary conditions are then related to the transports in the adjacent SCR or QNR. The regional analyses are one-dimensional, but multidimensional effects can be accounted for physically and semi-numerically by using the one-dimensional model as a basis [Ver87]. P-type substrate region is neglected here, but note that its effects can be accounted for by using the regional-analysis methodology.

3.2.1 Base Analysis

The base-region analysis extends the work by de Graaff and Kloosterman [Gra85]. They analyically (with interpolation for intermediate injection) solved the carrier transport problem in the quasi-neutral base for a representative exponential doping profile, thereby physically accounting for the dependence of the transit time on built-in electric field and bias conditions. In their analysis, however, they first assumed an arbitrary (constant) QNB width, and then perturbed the solution by letting the QNB width be a function of bias to empirically account for base-width modulation. This, in

effect, unrealistically portrays the doping profile as being dependent on bias voltages since the expression for it involves the arbitrary QNB width. To reduce the empiricism in accounting for base-width modulation, the doping profile is defined as in [Gra85], but the moving boundaries of the QNB are explicitly accounted for throughout the analysis. Thus, the following doping profile is assumed:

$$N_A(x) = N_{AO} \exp\left(- \frac{\eta x}{W_{BM}} \right) \tag{3.1}$$

where W_{BM} is the metallurgical base width.

If the hole current is negligible (i.e., if the forward or inverse current gain is sufficiently greater than unity), then the base transport current can be expressed as [Gra85]

$$I_C = -qA\bar{D}_n \left(\frac{2n + N_A}{n + N_A} \frac{dn}{dx} + \frac{n}{n + N_A} \frac{dN_A}{dx} \right). \tag{3.2}$$

Except for hard saturation, (3.2) is valid for all modes of operation. Integrating (3.2) over the QNB width W_B, from x_E to x_C (see Fig. 3.1), analytical solutions , I_{CL} and I_{CH}, for low and high injection, respectively, are obtained:

$$I_{CL} = \frac{qA\bar{D}_n\eta}{W_{BM}} \left(\frac{fn_E - n_C}{f - 1} \right) \equiv I_{CLF} - I_{CLR} \tag{3.3}$$

$$I_{CH} = \frac{2qA\bar{D}_n\eta}{W_{BM}} \left(\frac{n_E - n_C}{\ln(f)} \right) \equiv I_{CHF} - I_{CHR} \qquad (3.4)$$

where $n_E \equiv n(x_E)$, $n_C \equiv n(x_C)$, and $f \equiv \exp(\eta W_B/W_{BM})$. For ease of interpolation for intermediate injection, I_{CL} and I_{CH} are decomposed into forward and reverse components as indicated in (3.3) and (3.4). Now the closed-form expression for the forward and reverse component of I_C, for all injection levels, is given by a weighted linear sum of the corresponding low- and high- injection solutions [Gra85]:

$$I_C = I_{CF} - I_{CR} = \frac{I_{CLF} + C_F \frac{n_E}{N_{AE}} I_{CHF}}{1 + C_F \frac{n_E}{N_{AE}}} - \frac{I_{CLR} + C_R \frac{n_C}{N_{AC}} I_{CHR}}{1 + C_R \frac{n_C}{N_{AC}}} \qquad (3.5)$$

where $N_{AE} \equiv N_A(x_E)$, $N_{AC} \equiv N_A(x_C)$, and C_F and C_R are coupling coefficients.

For the transient model, the stored charge in the base region is now considered. Q_{BB}, the total hole charge in the QNB, is evaluated from I_C in (3.5) by using the fundamental Moll-Ross relation [Mol56]:

$$Q_{BB} \equiv qA \int_{x_E}^{x_C} p \, dx$$

$$= qA \frac{qA\bar{D}_n n_i^2}{I_C} \left[\exp\left(\frac{V_{BE}}{V_T}\right) - \exp\left(\frac{V'_{BCO}}{V_T}\right) \right] \qquad (3.6)$$

where V'_{BCO} is the quasi-Fermi potential separation at $x = x_C$.

This characterization is not only physically sound, but it is computationally efficient and it eliminates the need for another interpolation for the charge as was done in [Gra85]. (Regions of operation where (3.5) and (3.6) are invalid, i.e., hard saturation, are narrow, and are accounted for by extrapolation.) Note that the junction capacitances and diffusion capacitances are naturally accounted for in the time-derivative of Q_{BB} in (3.6) without having to introduce empirical transit times to relate charge and current as in the Gummel-Poon model. The decomposition and partitioning of Q_{BB} needed to properly represent the base charge dynamics, as well as other charging currents, are defined in Subsection 3.2.4.

3.2.2 Collector Analysis

The quasi-saturation model developed in Chapter 2 cannot completely characterize the collector transport because the space-charge region at the base-collector junction was neglected. To facilitate proper connections among different regions of operation, it is necessary to include the junction

space-charge region in the collector model. In the analysis of this region, the key equation is the expression for its width (on the collector side) W_{SCC}. This expression yields criteria for checking the operating conditions in the collector and provides a basis for linking the base and collector analyses.

With the assumption that the integral of the electric field across the entire epitaxial collector region equals the junction built-in potential ϕ_C plus the reverse-bias voltage $-V_{BC}$, W_{SCC} can be expressed in terms of I_C by applying Poisson's equation to the junction space-charge region:

$$W_{SCC} = \left[\frac{2\varepsilon A v_S \left(\phi_C - V_{BC} - I_C R_{EPI}\right)}{\left(I_O - I_C\right)}\right]^{1/2}. \qquad (3.7)$$

It must be noted that (3.7) is valid only for $I_C < I_O$ and $\phi_C > V_{BC}$. For $I_C > I_O$, a current-induced SCR forms away from the base-collector junction and invalidates (3.7). For $\phi_C < V_{BC}$, the base-collector junction becomes heavily forward-biased and the entire epitaxial collector is quasi-neutral.

When $I_C < I_O$, W_{SCC} can either expand or shrink with increasing I_C as can be understood from the two counteracting terms involving I_C in (3.7). The term in the numerator accounts for ohmic drop in the quasi-neutral region, whereas the term in the denominator represents the effect of space-

charge modulation by current-carrying electrons. The dominance of one effect over the other depends on the value of V_{BC} [Bow73]. To derive this critical value of V_{BC}, (3.7) is differentiated with respect to I_C to yield

$$\frac{dW_{SCC}}{dI_C} = \frac{\varepsilon v_S A(\phi_C - V_{BC} - I_O R_{EPI})}{\sqrt{2\varepsilon v_S A(\phi_C - V_{BC} - I_C R_{EPI})(I_O - I_C)^3}}. \qquad (3.8)$$

Thus, (3.8) provides the criteria for determining the behavior of W_{SCC} with respect to I_C for given V_{BC}. W_{SCC} expands with I_C if $V_{BC} \leq V_{CRIT} \equiv \phi_C - I_O R_{EPI}$ and contracts with I_C otherwise.

Now (3.7) and (3.8) can be used to derive the condition required for the collapse of the junction space-charge region, as well as the condition for the space-charging of the entire epitaxial collector. To summarize the results,

$$W_{SCC} = 0 \qquad \text{for } I_C \geq I_{SCC} \qquad (3.9)$$

$$\text{where } I_{SCC} \equiv \frac{\phi_C - V_{BC}}{R_{EPI}}, \text{ and} \qquad (3.10)$$

$$W_{SCC} = W_{EPI} \qquad \text{for } I_C < I_{SCCWEPI} \text{ and } V_{BC} \geq V_{CRIT} \text{ or} \qquad (3.11)$$

$$\text{for } I_C > I_{SCCWEPI} \text{ and } V_{BC} \leq V_{CRIT} \qquad (3.12)$$

$$\text{where } I_{SCCWEPI} \equiv \frac{I_O - \dfrac{2\varepsilon A v_S (\phi_C - V_{BC})}{W_{EPI}^2}}{1 - \dfrac{2\varepsilon A v_S R_{EPI}}{W_{EPI}^2}} . \tag{3.13}$$

There is another mechanism via which the entire epitaxial collector can become space-charged: the expansion of current-induced SCR described in Section 2.3. To derive the onset condition for the space-charging of the entire epitaxial collector region for this case, (2.6) can be used. From (2.6) and (2.2), it can be argued that W_{SCR} decreases with I_C for given V_{BC}. Thus, the following criteria can be derived:

$$W_{SCR} = W_{EPI} \quad \text{for } I_C \leq I_{SCRWEPI} \tag{3.14}$$

$$\text{where } I_{SCRWEPI} = I_O + \frac{2\varepsilon v_S A (\phi_C - V_{BC} - E_S W_{EPI})}{W_{EPI}^2} . \tag{3.15}$$

Equations (3.9) through (3.15) can be used to define different operating conditions as distinguished by their charge conditions. Fig. 3.2 shows the resulting four possible charge conditions. Note that these conditions are determined by the value of V_{BC} and I_C. This recognition is an important

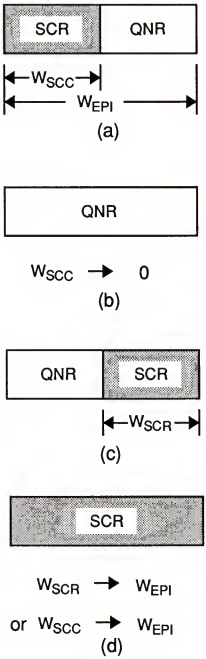


Fig. 3.2 Four possible charge conditions in the epitaxial collector.

consideration in developing the algorithm for the solution of model equation, as will be discussed in detail in Appendix A.

The stored charge in the QNRs for case (b) and (c) in Fig. 3.2 can be evaluated by using [Kul85] and (2.11), respectively. Note that Q_{QNR} is the neutralized excess charge stored in the quasi-neutral epitaxial collector region. Stored charge in the SCRs must also be evaluated not only to define the transient model, but also to physically link the base and collector transports based on the charge-neutrality, as will be described in Subsection 3.2.3. The stored charge in the space-charge region can analytically be derived based on the model assumptions in Section 2.3 prior to (2.3). To generalize the treatment, Q_{SCR} in (2.10) is now renamed Q_{EPI} to represent the charge stored in either the junction SCR or the current-induced SCR:

$$Q_{EPI} \equiv qA \int_{SCR} (N_{EPI} - \frac{I_C}{qAv_S}) dx \qquad (3.16)$$

$$= - qAW_{SCR}(\frac{I_C}{qAv_S} - N_{EPI}) \qquad \text{for } W_{SCR} \neq 0$$

$$= qAW_{SCC}(N_{EPI} - \frac{I_C}{qAv_S}) \qquad \text{for } W_{SCC} \neq 0.$$

This generalization is done to facilitate the linkage between the base and collector transports via total charge-

neutrality in p-n-n^+ region as will be described in Subsection 3.2.3. When the total charge neutrality is imposed in the p-n-n^+ region, the charge stored in the buried layer, Q_{BL}, must also be taken into account since a substantial electric field can exist at the epi-buried-layer junction. By Gauss's law,

$$Q_{BL} = \varepsilon E_{BL} A \tag{3.17}$$

where ε is the permittivity of silicon, and E_{BL} is the magnitude of the electric field at the epi-buried-layer junction. These regional charge components are used in Subsection 3.2.4 to define the corresponding charging currents for the charge-based BJT model.

3.2.3 Boundary Conditions and Linkages

The analysis in Subsection 3.2.1 formulated the QNB transport as a moving-boundary problem. Consequently its solutions are explicit functions of boundary locations (x_E and x_C) as well as carrier densities (n_E and n_C) at the boundaries:

$$I_C^B = f(x_E, n_E, x_C, n_C). \tag{3.18}$$

The boundary conditions at $x = x_E$ can be defined straightforwardly via the emitter-base junction SCR analysis. On the other hand, the boundary conditions at $x = x_C$ are complexly dependent on the collector transport, and hence link the base and collector transports.

First, the boundary conditions at $x = x_E$ are described. The carrier density is evaluated from the pn product by using quasi-neutrality:

$$n_E = \frac{N_{AE}}{2} \{ [1 + (\frac{2n_i}{N_{AE}})^2 \exp(\frac{V_{BE}}{V_T})]^{1/2} - 1\}. \qquad (3.19)$$

The boundary location x_E is the metallurgical junction location (x_{JE}) plus the width of the space-charge region on the base side of the base-emitter junction. This width is semi-physically modeled [Lio87] as follows:

$$W_{SE} = W_{SEO} [\frac{\phi_E - V_{BE}}{\phi_E}]^{m_E} \quad \text{for } V_{BE} \le \phi_E - \Delta\phi_E \qquad (3.20)$$

$$= W'_{SEO} \exp (- \frac{V_{BE}}{4V_T}) \quad \text{for } V_{BE} > \phi_E - \Delta\phi_E \qquad (3.21)$$

where W_{SEO} is the base-side width of the base-emitter junction SCR at equilibrium, and W'_{SEO} and $\Delta\phi_E$ are related to W_{SEO} and m_E by imposing the continuities of W_{SE} and its derivative with respect to V_{BE} at $V_{BE} = \phi_E - \Delta\phi_E$.

To evaluate the boundary conditions at $x = x_C$, the collector operating conditions discussed previously must be carefully examined to establish physical linkages between the base and collector transports. As far as physical linkages are concerned, the four different operating conditions illustrated in Fig. 3.2 can be divided into two distinct operating modes based on the nature of interaction with the base transport.

For one mode, the presence of a space-charge region at the base-collector junction causes the modulation of x_C (Early effect) to be the dominant mechanism coupling the base and collector transports. This coupling can physically be modeled by ensuring total charge-neutrality across the entire base-collector (p-n-n^+) junction.

For the other mode, the space-charge region at the base-collector junction virtually vanishes due to heavy (internal) forward biasing of the junction, and the modulation of n_C (quasi-saturation) is the dominant mechanism coupling the base and collector transports. For this case, the quasi-Fermi potential separation at the base-collector junction becomes a key physical variable in the analysis because it can be used to describe the transport in the quasi-neutral portion of the epitaxial collector as well as n_C in (3.18).

For cases (b) and (c), the junction SCR vanishes and V_{BCO} merges with V'_{BCO} defined in connection with (3.6): $n_C(V_{BCO})$ is defined analogously to (3.19) and $x_C = x_{JC}$, where x_{JC} denotes the location of the base-collector metallurgical junction. For cases (a) and (d), x_C is defined by ensuring neutrality of the total charge across the entire base-collector (p-n-n^+) junction, and $n_C \approx I_C/qAv_S$ assuming that the electron velocity saturates in the high-field space-charge region.

Note that the collector model describes x_C and n_C in terms of I_C and V_{BC} for all operating modes. These collector-side QNB boundary conditions, in conjunction with (3.5), define the system of model equations. The basic model algorithm, which requires numerical solution as discussed in Section 5.3, is flowcharted in Fig. 3.3. Solution of this algorithm yields I_C and enables calculations of the regional charges. Detailed model algorithms are discussed in Appendix A.

As indicated in Fig. 3.4, the model is highly physical and properly simulates all possible regions of operation. (Note from the collector and base analyses that hard saturation is accounted for only empirically.) To illustrate this generality, note that case I represents a transition through saturation, ohmic quasi-saturation, ohmic quasi-

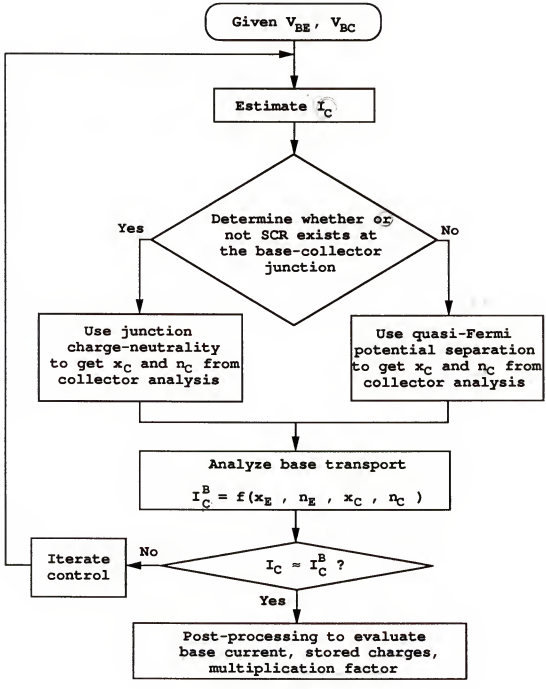

Given V_{BE}, V_{BC}
Estimate I_C
Determine whether or not SCR exists at the base-collector junction
Yes
No
Use junction charge-neutrality to get x_C and n_C from collector analysis
Use quasi-Fermi potential separation to get x_C and n_C from collector analysis
Analyze base transport
$I_C^B = f(x_E , n_E , x_C , n_C)$
Iterate control
No
$I_C \approx I_C^B$?
Yes
Post-processing to evaluate base current, stored charges, multiplication factor

Fig. 3.3 A flowchart of the BJT model algorithm

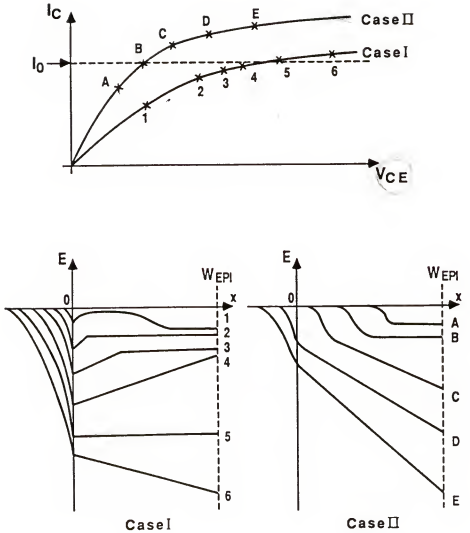


Fig. 3.4 Electric field profiles for designated points on I_C versus V_{CE} (cases I and II). The two cases illustrate the two distinct mechanisms leading to the complete space-charging of the entire epitaxial collector region.

saturation, and a forward-active mode in which the base-collector junction space-charge region extends throughout the epitaxial collector. It is interesting to note that, in general, the forward-active mode cannot be unambiguously defined in the model, which reflects its physical nature. The model predicts different forward-active-mode characteristics without having to explicitly define that mode of operation. For example, the model can account for possibly different characterizations of base-width modulation depending on the charge condition of the collector.

3.2.4 Charging Currents

In the preceding regional analyses in Section 2.3 and Subsections 3.2.1 and 3.2.2, expressions for regional stored charges have been derived. Derivatives of these charges can define the charging currents in the large-signal charge-based BJT model. However, some of the charge components must be partitioned in order to account for their distributed nature.

First, the charge stored in the base region is considered. The hole charge Q_{BB} in the QNB can, based on quasi-neutrality, be expressed in terms of electron charge and ionized dopant charge:

$$Q_{BB} = qA\int_{x_E}^{x_C} n\,dx + qA\int_{0}^{W_{BM}} N_A\,dx - qA\int_{0}^{x_E} N_A\,dx - qA\int_{x_C}^{W_{BM}} N_A\,dx \quad (3.22)$$

$$\equiv Q_N + Q_{BO} + Q_{JE} + Q_{JC} \quad (3.23)$$

where, using (3.1), depletion charges, Q_{JE} and Q_{JC}, can be evaluated as

$$Q_{JE} = -\frac{qAN_{AO}W_{BM}}{\eta}\,[1 - \exp(-\frac{\eta x_E}{W_{BM}})] \quad (3.24)$$

and

$$Q_{JC} = -\frac{qAN_{AO}W_{BM}}{\eta}\,[\exp(-\frac{\eta x_C}{W_{BM}}) - \exp(-\eta)]. \quad (3.25)$$

Note, from the charge-neutrality, that Q_{JC} can be related to charge components stored in the collector region, as defined in (3.16) and (3.17):

$$Q_{JC} + Q_{EPI} + Q_{BL} = 0\ . \quad (3.26)$$

In the charge-based model, charging currents are defined by the time-derivatives of the associated charges. Thus, the total charging current for the base region is

$$\frac{dQ_{BB}}{dt} = \frac{dQ_N}{dt} + \frac{dQ_{JE}}{dt} + \frac{dQ_{JC}}{dt}\ . \quad (3.27)$$

Note that the constant term Q_{BO} is dropped in (3.27).

Physically, Q_N (the electron charge in the QNB) in (3.23) is shared between the emitter and collector terminals for transient operation [Fos86]. This sharing can be simulated by partitioning Q_N:

$$Q_N = Q_N^E + Q_N^C \approx Q_N f_B + Q_N (1 - f_B) \tag{3.28}$$

where f_B is a partition factor assumed to be a constant between 0 and 1, although in reality it would be bias-dependent.

The junction depletion charge Q_{JE} can be attributed to the emitter. However, Q_{JC} must be partitioned between emitter and the collector. This is because the emitter terminal, which is the main source of electron flow, can contribute to the charging of the base-collector junction space-charge region. It is uncertain how Q_{JC} can physically be partitioned in general. However, a case exists where there is no ambiguity as to the physical method of partitioning. Such a case is found when a current-induced space-charge region forms in the epitaxial collector as indicated by case (c) in Fig. 3.2. In this case, charging does not involve any hole current. It is obvious that Q_{EPI} can be attributed to the emitter, while Q_{BL} can be attributed to the collector. Here,

the partitioning for this case is assumed to be applicable to other cases without significant error, rather than choosing a totally arbitrary partitioning. Then, from (3.26) and (3.27), the total charging current for the base region can be partitioned as follows:

$$\frac{dQ_{BB}^{E}}{dt} = \frac{dQ_{N}^{E}}{dt} + \frac{dQ_{JE}}{dt} - \frac{dQ_{EPI}}{dt} \tag{3.29}$$

and

$$\frac{dQ_{BB}^{C}}{dt} = \frac{dQ_{N}^{C}}{dt} - \frac{dQ_{BL}}{dt} \; . \tag{3.30}$$

The stored (neutralized) charge in the quasi-neutral epitaxial collector region can also be partitioned in the same manner as Q_N is partitioned:

$$Q_{QNR} = Q_{QNR}^{E} + Q_{QNR}^{C} \approx Q_{QNR} f_C + Q_{QNR} (1 - f_C) \tag{3.31}$$

where f_C is another constant partition factor.

3.2.5 Model Network Representation

To complete the model, the base current, the hole charge stored in the emitter, and the parasitic series resistances

must be included. (The charging currents associated with mobile charges stored in the junction space-charge regions [Lio87] can be assumed to be negligible relative to the other charging currents.)

The base current I_B comprises four components: back injection into the emitter, base-emitter junction space-charge region recombination, recombination in the QNB, and the recombination in the epitaxial collector. Generally then, I_B can be expressed as follows:

$$I_B = I_{EO}[\exp(\frac{V_{BE}}{V_T}) - 1] + I_{SEO}[\exp(\frac{V_{BE}}{n_{EB}V_T}) - 1] + \frac{Q_N}{\tau_B} + \frac{Q_{QNR}}{\tau_C} \quad (3.32)$$

where all the parameters introduced are constants. The latter two current components are typically negligible in advanced BJTs, but when they are significant, they are represented well by constant recombination lifetimes τ_B and τ_C. The first component in (3.32) relates directly to the hole charge Q_{EE} stored in the emitter. In general [Yun88]

$$Q_{EE} = \tau_E I_{EO}[\exp(\frac{V_{BE}}{V_T}) - 1] \quad (3.33)$$

where τ_E, which governs hole transport and recombination in both the monocrystalline and polycrystalline portions of the emitter, is constant and independent of V_{BE} and V_{BC} because of

prevalent low-injection conditions in the highly doped emitter.

The network representation of the charge-based BJT model is shown in Fig. 3.5. Note that the charge storage associated with the collector-substrate junction is accounted for by dQ_{SUB}/dt in the network representation, even though the p-substrate has not been explicitly considered in the model. Note also that the charging currents are simulated directly without introducing capacitors. For all the charges (except Q_{SUB}), the associated charging currents are

$$\frac{dQ}{dt} = \frac{\partial Q}{\partial V_{BE}} \frac{dV_{BE}}{dt} + \frac{\partial Q}{\partial V_{BC}} \frac{dV_{BC}}{dt} \tag{3.34}$$

where the (trans)capacitance coefficients are evaluated from the steady-state analyses. This charge-based formalism is more physical and representative of the distributed charge dynamics than is the conventional capacitance-based model such as Gummel-Poon. Actual nonreciprocal trans-capacitances for the BJT are properly simulated, thereby allowing a first-order accounting for possibly significant non-quasi-static behavior [Fos86].

Now all the charging current components in parallel can be lumped together to define two current sources:

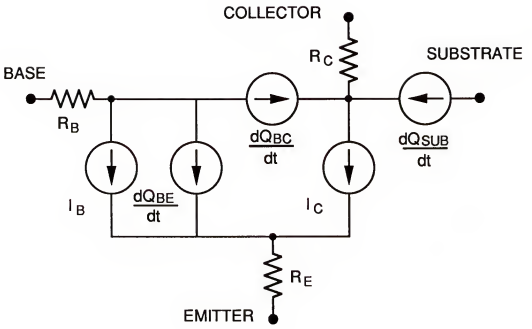


Fig. 3.5 Network representation of the charge-based bipolar n-p-n transistor model

$$\frac{dQ_{BE}}{dt} = \frac{dQ_N^E}{dt} + \frac{dQ_{JE}}{dt} - \frac{dQ_{EPI}}{dt} + \frac{dQ_{QNR}^E}{dt} + \frac{dQ_{EE}}{dt} \tag{3.35}$$

and

$$\frac{dQ_{BC}}{dt} = \frac{dQ_N^C}{dt} - \frac{dQ_{BL}}{dt} + \frac{dQ_{QNR}^C}{dt} . \tag{3.36}$$

In this dissertation, V_{BE} and V_{BC} denote internal voltages, viz., the terminal voltages minus ohmic drops in the extrinsic regions of the device. These drops are accounted for by the series resistances R_B, R_E, and R_C. The emitter resistance R_E is a constant as is the extrinsic collector resistance R_C. Note that the nonconstant resistance of the intrinsic epitaxial collector is accounted for rigorously in the implicit characterization of I_C described previously. The base resistance comprises a constant extrinsic component R_{Bx} and a nonconstant intrinsic component R_{Bi}:

$$R_B = R_{Bx} + R_{Bi} . \tag{3.37}$$

To properly characterize R_{Bi}, which is affected by conductivity modulation both in the quasi-neutral base and extended base (i.e., the QNR in the collector region),

emitter debiasing [Ver87] must be accounted for. In this one-dimensional BJT model, however, debiasing is not considered. Hence only a one-dimensional description of the R_{Bi} modulation is given, which can be the basis for a complete multidimensional characterization.

This modulation is implicit in the model characterization of Q_{BB} and Q_{QNR} described previously. Thus, straightforward charge-control accounting for lateral hole flow in the (widened) base yields the bias-dependent R_{Bi}:

$$R_{Bi} \approx C_{RBi} \frac{W_E^2}{(Q_{BB} + Q_{QNR})} \qquad (3.38)$$

where $C_{RBi} \equiv \frac{f_E}{4\bar{\mu}_p}$, W_E is the emitter width, $\bar{D}_n$ is an average hole mobility, and f_E is a geometrical constant that is less than unity for finite emitter length.

3.3 Experimental Support

Test devices representative of the advanced bipolar technology were used to verify the model and to define a parameter-extraction algorithm. The devices were fabricated at Motorola, Inc., using the MOSAIC III [Zde87] process (the third-generation Motorola Oxide Isolation Self Aligned

Implanted Circuits Process). In MOSAIC III technology, a double-polysilicon process, in conjunction with a "boot-shaped" sidewall spacer technique, enables a self-aligned submicrometer emitter structure with a lithography greater than 1 μm. This fabrication process yields improved speed-power performance and integration density due to reductions in the ratio of the extrinsic to intrinsic device area and in the extrinsic base resistance. The BJTs produced thus are useful in supporting the one-dimensional model, and also in identifying significant multidimensional effects that must be added to the model.

3.3.1. Parameter Extraction

There are two different kinds of parameters. Device parameters are related to the layout and thus can differ from device to device. The model has three device parameters, as tabulated in Table 3.1. Model parameters are process-dependent and thus usually common to all the BJT devices on the same chip. There are 31 model parameters, as tabulated in Tables 3.2, 3.3, and 3.4 for three different groups to be discussed below.

Parameter extraction for the model is straightforward due to its highly physical nature. Some of the parameters can be evaluated directly from layout and process information.

TABLE 3.1
DEVICE PARAMETERS FOR CHARGE-BASED BJT MODEL

Name	Description	Unit
AE	Effective Emitter Area	m^2
WE	Effective Emitter Width	m
AC	Effective Collector Area	m^2

TABLE 3.2
PHYSICAL MODEL PARAMETERS FOR CHARGE-BASED BJT MODEL

Name	Description	Units
UNEPI	Electron mobility in epitaxial collector	$cm^2\ V^{-1}\ s^{-1}$
NEPI	Epitaxial doping density	cm^{-3}
WEPI	Epitaxial collector thickness	m
WBM	Metallurgical base width	m
NAO	Extrapolated peak base doping density at emitter-base junction	cm^{-3}
ETA	Base doping gradient factor	-
VS	Electron saturation velocity	$cm\ s^{-1}$
AI	Pre-exponential constant for impact-ionization coefficient	cm^{-1}
BI	Exponential constant for impact-ionization coefficient	$V\ cm^{-1}$

TABLE 3.3
SEMI-PHYSICAL MODEL PARAMETERS FOR CHARGE-BASED BJT MODEL

Name	Description	Units
JEO	Emitter saturation current density	A m^{-2}
JSEO	Base-emitter SCR saturation current density	A m^{-2}
NEB	Base-emitter SCR emission coefficient	-
WSEO	Base-emitter equilibrium SCR width in base	m
PE	Base-emitter junction potential	V
ME	Base-emitter junction exponential factor	-
PC	Base-collector junction potential	V
DNB	Effective electron diffusivity in base	cm^2 s^{-1}
CJS	Zero-bias collector-substrate capacitance	F m^{-2}
PS	Substrate junction potential	V
MS	Substrate junction exponential factor	-
CRBI	Intrinsic base resistance coefficient	V s cm^{-2}
TB	carrier lifetime in base	s
TC	carrier lifetime in collector	s

TABLE 3.4
EMPIRICAL MODEL PARAMETERS FOR CHARGE-BASED BJT MODEL

Name	Description	Units
CIF	Forward current coupling coefficient	-
CIR	Reverse current coupling coefficient	-
RC	Extrinsic base resistance	Ω
RB	Extrinsic base resistance	Ω
RE	Extrinsic emitter resistance	Ω
FB	Base-charge partition factor	-
FC	Collector-charge partition factor	-
TE	Emitter hole transit time	s

Other parameters that are closely related to the device physics can be tightly constrained in value in the optimization process. The remaining few parameters are empirically defined, but can be optimized without much rigor. To facilitate the discussion, the set of model parameters is categorized as follows into three groups according to their nature as discussed above.

(1) Physical parameters: N_{AO}, W_{BM}, η(≡ETA), W_{EPI}, N_{EPI}, μ_n(≡UNEPI), v_S, A_i, B_i (The last two define the impact ionization coefficient, as will be described in Chapter 4).

(2) Semi-physical parameters: J_{SEO}, J_{EO}, n_{EB}, ϕ_E(≡PE), m_E, W_{SEO}, ϕ_C(≡PC), $\bar{D}_n$(≡DNB), C_{RBi}, τ_B(≡TB), τ_C(≡TC), ϕ_S(≡PS), m_S, C_{JS}.

(3) Empirical parameters: C_F(≡CIF), C_R(≡CIR), f_B, f_C, τ_E(≡TE), R_B, R_C, R_E.

The parameter A_C, which is larger than A_E, was added to the model to account for collector spreading, viz., the lateral diffusion of carriers in the collector under quasi-saturation conditions. This was done because the preliminary verification/parameter extraction revealed that collector spreading in the small devices significantly ameliorates the one-dimensional quasi-saturation effects, which are physically modeled. The use of a constant A_C is only a first-

order accounting; a more rigorous modification of the model may be needed for a multidimensional version of the model.

Now a general scheme for extracting the model parameters from given geometrical, process, and sample measured data is discussed. First, the physical parameters are assigned default values derived from the available geometrical (mask layout) and process information. N_{AO}, W_{BM}, η, N_{EPI}, and W_{EPI} are estimated from the drawn emitter area and knowledge of the emitter sidewall spacer dimension [Hwa87]. A_E can be significantly smaller than the drawn emitter area in the MOSAIC devices because of the spacer. For our test devices, the estimated width of the spacer is 0.4 μm.

Next, the semi-physical parameters, which may be coupled to physical parameters, are extracted via mini-optimizations based heavily on measured β versus V_{BE} characteristics. The extraction scheme is outlined as follows.

(1) Measure β versus V_{BE} for $V_{BC} = 0$ and $V_{BC} < 0$; also generate $I_C(V_{BE})$ (Gummel plot).

(2) By extrapolation of $I_C(V_{BE})$, estimate the saturation current I_S.

(3) From the measured β_{MAX} and I_S, estimate I_{EO}, neglecting Q_N/τ_B. J_{SEO} and n_{EB} are then extracted (optimized) from the β falloff at low V_{BE}.

(4) From the gradual β falloff at moderate V_{BE} (emitter-side base-width modulation) for $V_{BC} = 0$, extract (optimize) ϕ_E, m_E, W_{SEO}.

(5) From β versus V_{BC} for moderate V_{BE}, extract (optimize) ϕ_C.

(6) $\bar{D}_n$ can now be calculated from $I_S \approx qA\bar{D}_n n_i^2 \eta / \{[1-\exp(-\eta)]N_{AO}W_{BM}\}$, which is valid at low V_{BE} as implied by (3.3).

(7) From the β falloff at high V_{BE}, extract (optimize) τ_C and A_C.

Finally, the empirical parameters are extracted by optimization. Initial values for R_B, R_E, and R_C can be estimated by standard methods [Get78], but to extract representative values compatible with the physical model, subsequent optimization is needed. Note that R_C in the model is the extrinsic collector resistance and hence must be extracted in the saturation region of the transistor. The optimization, which fits measured current-voltage characteristics, also evaluates C_F and C_R (predominantly from the high-V_{BE} region). This optimization is done iteratively in conjunction with the mini-optimizations used to extract the semi-physical parameters.

The parameters f_B, f_C, and τ_E are significant only for transient simulations. They therefore cannot be extracted from dc measurements, but can be derived, for example, from high-frequency s-parameter measurements [Fos87].

Parameter optimization needed as part of parameter extraction is a tedious process when performed without the aid of optimization software. To facilitate the parameter optimization, a simple optimization program was written, as is described in Appendix B.

3.3.2. Simulations

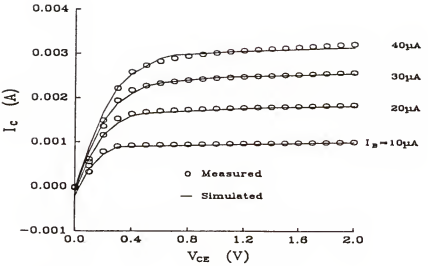
Here, SPICE simulations of the test devices are exemplified for model verifications. Simulations were done with MMSPICE which was created by writing the charge-based BJT model into SPICE2 source code, as will be described in Chapter 5. Transistors with two different emitter areas were simulated. The larger BJT, with a drawn emitter area of 2 X 4 μm^2, was first simulated, doing the parameter extraction straightforwardly as described in Subsection 3.3.1. Values of extracted parameters were close to the expected values inferred from the work [Zde87] and further optimization by the parameter optimizer described in Appendix B did not result in any significant departures in parameter values. The BJT, with a drawn emitter area of 1.5 X 4 μm^2, was, then, simulated with reasonably good accuracy simply by scaling down A_E and A_C.

Simulated $I_C(V_{CE}, I_B)$ and $\beta(V_{BE}, V_{BC})$ curves are contrasted to the corresponding measurements in Figs. 3.6 and

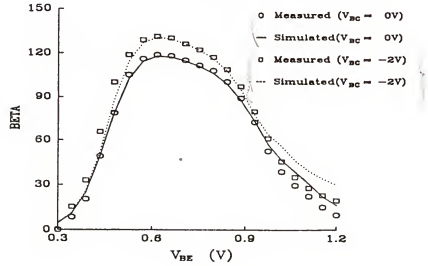
3.7 for the two devices. For the larger BJT in Fig. 3.6, the simulations were excellent. The discrepancy between the measured and predicted β in Fig. 3.6(b) at high V_{BE} is beyond the normal operating range. The slight discrepancy at low V_{BE} in Fig. 3.6(b) for $V_{BC} = -2V$ is thought to be due to impact ionization that occurs in the epitaxial collector region. A much more profound impact ionization effect was observed as V_{BC} becomes more negative and the electric field in the collector becomes higher. As will be described in Chapter 4, impact ionization can be physically accounted for in the model because of its inherent general representation of the electric field in the collector.

For the smaller device in Fig. 3.7, the simulations were reasonably good. The discrepancy between the measured and predicted β in Fig. 3.7(b) is attributed predominantly to lateral injection at the emitter-base junction, which is not simulated, but could be accounted for by simply adding another component to I_B in (3.32) [Ver87]. Note that collector spreading is significant at high currents in these devices as indicated by the extracted A_C being greater than A_E. This spreading ameliorates the quasi-saturation effects, and hence influences the V_{BC} dependences of I_C and β. The slight discrepancies in Fig. 3.7(a) between the measured and predicted I_C for high I_B may thus indicate a need for a more

Fig. 3.6 Measured and simulated $I_C(V_{CE}, I_B)$ in (a) and $\beta(V_{BE}, V_{BC})$ in (b) for the test BJT with drawn emitter area of 2 X 4 μm^2.

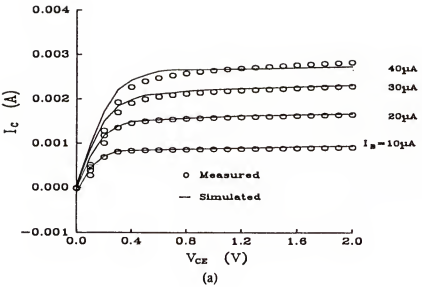

0.004
0.003
0.002
0.001
0.000
−0.001
I_C (A)
0.0
0.4
0.8
1.2
1.6
2.0
V_{CE} (V)
40μA
30μA
20μA
I_B=10μA
Measured
Simulated

(a)

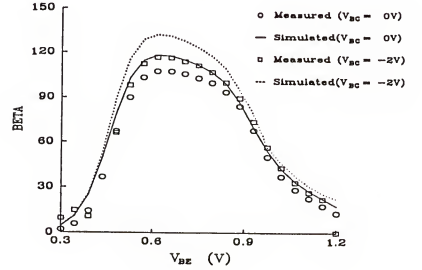

150
120
90
60
30
0
BETA
0.3
0.6
0.9
1.2
V_{BE} (V)
Measured (V_{BC} = 0V)
Simulated(V_{BC} = 0V)
Measured (V_{BC} = −2V)
Simulated(V_{BC} = −2V)

(b)

Fig. 3.7 Measured and simulated $I_C(V_{CE}, I_B)$ in (a) and $\beta(V_{BE}, V_{BC})$ in (b) for the test BJT with drawn emitter area of 1.5 X 4 μm^2. These simulations involved no parameter extraction beyond that done for the larger-device simulations in Fig. 3.6.


0.004
0.003
0.002
0.001
0.000
−0.001
I_C (A)
0.0
0.4
0.8
1.2
1.6
2.0
V_{CE} (V)
40μA
30μA
20μA
I_B=10μA
Measured
Simulated

(a)


150
120
90
60
30
0
BETA
0.3
0.6
0.9
1.2
V_{BE} (V)
Measured (V_{BC} = 0V)
Simulated(V_{BC} = 0V)
Measured (V_{BC} = −2V)
Simulated(V_{BC} = −2V)

(b)

physical accounting for collector spreading. Since the smaller device simulations were done without any further parameter extraction, the model seems to be scalable with reasonable accuracy down to the 1.5 μm drawn emitter width, provided first-order accountings for the multidimensional currents are included.

CHAPTER 4
PHYSICAL MODELING OF WEAK IMPACT IONIZATION

4.1 Introduction

High electric fields that can exist in the epitaxial collector region can cause carriers to gain sufficient kinetic energy to generate electron-hole pairs via impact ionization as they traverse the high-field region, as illustrated in Fig. 4.1. In n-p-n devices the generated electrons add to the collector current and generated holes are swept into the base region to drive the transistor. For sufficient impact ionization, the base current can change its direction if the base-emitter junction is driven by a voltage source.

In the advanced BJTs, the epitaxial collector doping density tends to be high (typically > 10^{16} cm^{-3}) to suppress quasi-saturation or base push-out that can occur due to increased current density in scaled devices. The high epitaxial doping density results in high peak electric field at the base-collector junction and can cause significant impact ionization. However, at high currents, quasi-saturation effects reduce the peak electric field and eventually relocate the point of peak electric field from the

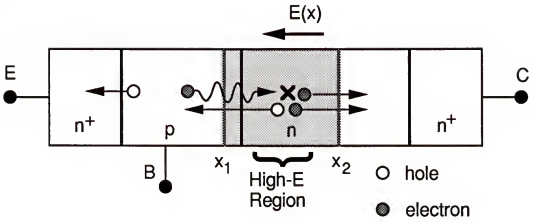


Fig. 4.1 Impact ionization mechanism in the bipolar transistor.

base-collector junction to the buried-layer high-low junction.

The conventional methods of empirically modeling impact ionization are inadequate because the multiplication factor due to impact ionization is determined by the electric field distribution that is complexly dependent on both V_{BE} and V_{BC}. The physical modeling of impact ionization should account for all possible operating conditions in the epitaxial collector as described in Section 3.2.2. This modeling is of interest for several reasons. First, impact ionization can be a serious issue for future bipolar device scaling. Second, the reversal of base current can be used for novel circuit applications [Sak89]. Thirdly, the analysis provides some indirect experimental support for the collector model, e.g., confirmation of the high-field-region relocation at high currents.

4.2 Physical Model

The model is restricted to the weak impact ionization which is amenable to simple analysis and is the most important impact ionization mechanism in modeling advanced BJTs. For weak impact ionization, the multiplication factor is expressed as

$$M = 1 + \int_{EPI} \alpha_n(x)\ dx \tag{4.1}$$

$$= 1 + \int_{x_1}^{x_2} \alpha_n(x)\ dx + \int_{UNMOD} \alpha_n(x)\ dx \tag{4.2}$$

where α_n is the ionization coefficient for majority electrons, x_1 and x_2 are bias-dependent boundaries of the high-field space-charge region, and UNMOD denotes the unmodulated quasi-neutral region where the impact ionization is small, but may have to be considered, as is discussed later. The impact ionization coefficient is, based on a local "lucky electron" model [Sho61], given by

$$\alpha_n(x) = A_i \exp\left(- \frac{B_i}{E(x)} \right). \tag{4.3}$$

Impact ionization occurs predominantly in the high-field space-charge region in the collector. The impact ionization in the conductivity-modulated quasi-neutral region is negligible. However the impact ionization arising in the unmodulated quasi-neutral region can become substantial enough to cause noticeable discontinuity in the simulated M when it is neglected. So M is evaluated by summing up the contributions due to both the space-charge region and unmodulated QNR. Evaluating the impact ionization occurring

in the unmodulated quasi-neutral region is straightforward because the field can be assumed to be constant.

In the space-charge region, E(x) can be approximated as being spatially linear based on the assumptions that majority-carrier current flows predominantly by drift at the saturation velocity v_S and that the minority-carrier density and current are negligible. This allows the spatial integration to be transformed into an integration in terms of electric field. Thus, (4.2) can be rewritten as follows:

$$M = 1 + \int_{E_L}^{E_R} \alpha_n(E) \left(\frac{dE}{dx}\right)^{-1} dE + W_{QNR} A_i \exp\left(-\frac{B_i}{E_{QNR}}\right) \qquad (4.4)$$

where, from Poisson's equation, dE/dx in the SCR is a constant given by $(q/\varepsilon)(N_{EPI} - I_C/qAv_S)$. The boundary values of E, E_L, and E_R, which are extracted from the BJT model analysis defined in Chapter 3, as well as W_{QNR} and E_{QNR}, the magnitude of electric field in the unmodulated region, are summarized below for four possible different charge conditions in the collector depicted in Fig. 3.2. These boundary conditions can be derived straightforwardly following the same method as was used to derive (3.7).

(1) For case (a),

$$E_R = \frac{I_C}{qA\mu_n N_{EPI}}$$

$$E_L = E_R + W_{SCC}\,\frac{I_O - I_C}{\varepsilon A v_S}$$

$$E_{QNR} = E_R$$

$$W_{QNR} = W_{EPI} - W_{SCC} \ .$$

(2) For case (b),

$$E_{QNR} = \frac{I_C}{qA\mu_n N_{EPI}}$$

$$W_{QNR} = \frac{\phi_C - V_{BC}}{E_{QNR}} \ .$$

(3) For case (c),

$$E_L = E_S$$

$$E_R = 2\,\frac{\phi_C - V_{BC}}{W_{SCR}} - E_S$$

$$E_{QNR} = 0 \ .$$

(4) For case (d),

$$E_L = \frac{\phi_C - V_{BC}}{W_{EPI}} + W_{EPI} \frac{I_Q - I_C}{2\varepsilon A v_S}$$

$$E_R = E_L - W_{EPI} \frac{I_Q - I_C}{\varepsilon A v_S}$$

$$W_{QNR} = 0 \ .$$

The integration in (4.4) is done numerically by using Romberg's method, which is a very efficient algorithm combining the trapezoidal method with Richardson extrapolation. The current generated by the impact ionization is $(M-1)I_C$; generated holes are swept into the base, intensifying the transistor action, and the electrons flow out the collector terminal. For weak impact ionization, the generated current is negligible, and the BJT model, in Fig. 3.5, is properly extended by augmenting the base current element as follows:

$$I_B \rightarrow I_B - (M-1)I_C$$

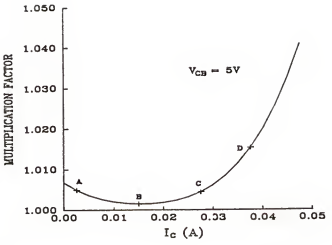
where I_C is implicitly characterized by the model in Chapter 3. The impact ionization can be turned on and off by the user.

4.3 Experimental Support

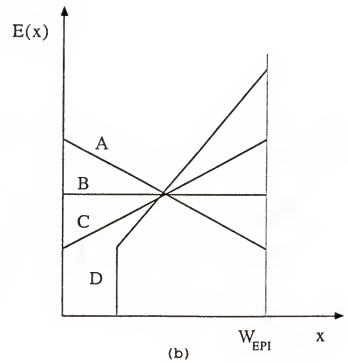
The impact ionization model was used to simulate advanced BJTs fabricated at Motorola, Inc. [Zde87]. In Fig. 4.2(a), a simulated M is shown as a function of I_C for V_{CB} set at 5V. To facilitate a qualitative understanding of the complicated dependence of M on I_C, the electric field distributions in the epitaxial collector corresponding to designated points on the curve in Fig. 4.2(a) are shown in Fig. 4.2(b). The initial decrease of M with increasing I_C is caused by the decrease in the peak electric field at the metallurgical collector-base junction. M reaches its minimum when I_C is approximately the velocity-saturated $I_O = qAN_{EPI}v_S$. Then, M increases with I_C as the peak electric field increases at the epi-buried layer (low-high) junction. The subsequent base widening initially increases M as E increases, but ultimately M should begin to decrease as α_n saturates at very high E and the width of the current-induced space-charge region shrinks. This decrease of M occurs at currents larger than those covered in Fig. 4.2.

The utility of the BJT model with the impact ionization physically accounted for is indicated in Figs. 4.3 and 4.4, where measured and simulated current-voltage characteristics are shown. In Fig. 4.3(a) measured I_C and I_B versus V_{BE} are

Fig. 4.2 Simulated multiplication factor versus I_C (a) and electric field distributions in the epitaxial collector (b) corresponding to designated points in (a)


1.050
1.040
1.030
1.020
1.010
1.000
MULTIPLICATION FACTOR
V_{CB} = 5V
A
B
C
D
0.00
0.01
0.02
0.03
0.04
0.05
I_C (A)

(a)

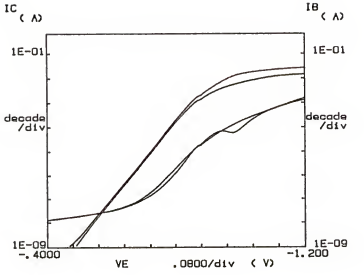

E(x)
A
B
C
D
W_{EPI}
x

(b)

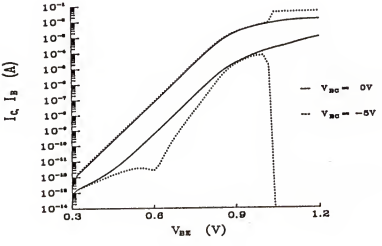
plotted for V_{CB} = 0 and 5V. The effects of impact ionization on the V_{CB} =5V plots are evident; I_B shows anomalous dips and I_C reveals avalanching at high V_{BE}. The corresponding simulations are plotted in Fig. 4.3(b). The experimental trends are predicted even though no model parameter optimization was done. For high V_{BE}, the simulated I_B becomes negative; experimentally this happened for V_{CB} slightly higher than 5V. In Fig. 4.4, measured and the simulated $I_C(V_{CE}, I_B)$ characteristics are shown. With qualitative agreement between the measurements and the simulations, these characteristics show the common-emitter breakdown (BV_{CEO}), including unusual negative-resistance properties. These anomalies and those in Fig. 4.3 can be attributed to the complexity of $M(V_{BE}, V_{CB})$ in the advanced BJT discussed with reference to Fig. 4.2.

The impact-ionization model that has been used here has evolved primarily from work involving silicon MOSFETs. To verify the model for bipolar transistors, M was experimentally characterized as follows. I_B and I_C are measured as a function of V_{BE} for two different values of V_{CB}: (I_{BR}, I_{CR}) for $V_{CB} = V_{CBR}$ and (I_{BM}, I_{CM}) for $V_{CB} = V_{CBM}$. The lower value of V_{CB}, V_{CBR}, serves as a reference in the sense that impact ionization is negligible. At the higher value of V_{CB}, I_C increases via two possible mechanisms: base-width modulation and/or impact ionization. The effects of these two

Fig. 4.3 Measured (a) and simulated (b) collector and base currents versus V_{BE} for V_{CB} = 0 and 5V.

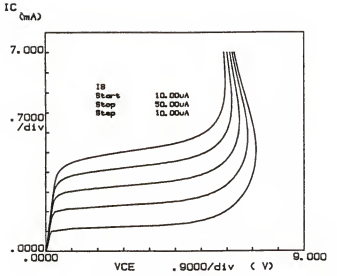

IC
(A)
IB
(A)
1E-01
1E-01
decade
/div
decade
/div
1E-09
1E-09
-.4000
VE
.0800/div
(V)
-1.200

(a)

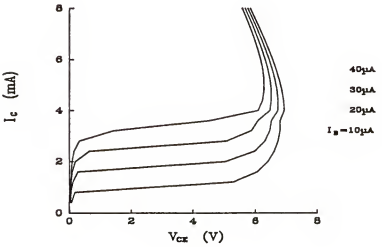

I_C, I_B (A)
V_{BE} (V)
0.3
0.6
0.9
1.2
V_{BC} = 0V
V_{BC} = −5V

(b)

Fig. 4.4 Measured (a) and simulated (b) collector currents versus V_{CE} and I_B.


IC (mA)
7.000
.7000 /div
.0000
.0000
IB
Start 10.00uA
Stop 50.00uA
Step 10.00uA
VCE .9000/div (V)
9.000

(a)


I_C (mA)
8
6
4
2
0
0
2
4
6
8
V_{CE} (V)
40μA
30μA
20μA
I_B=10μA

(b)

mechanisms can be separated if the effect of base-width modulation on the base current can be neglected, which is a valid assumption for thin base transistors. Thus,

$$I_{CM} = I_{CR} + \Delta I_{II} + \Delta I_{BWM} \tag{4.5}$$

and

$$I_{BM} = I_{BR} - \Delta I_{II} \ . \tag{4.6}$$

From (4.5) and (4.6), ΔI_{II} can be solved for. Then M can be experimentally estimated by using

$$M = \frac{I_{CM}}{I_{CM} - \Delta I_{II}} \ . \tag{4.7}$$

The measured M is plotted as a function of I_C in Fig. 4.5. The variation of M with I_C shows the same trends as the simulated $M(I_C)$ in Fig. 4.2, but there are two notable discrepancies. First, the minimum measured M is less than unity, which is physically impossible. This error is a result of anomalous behavior of the measured I_B which increased with V_{CB} for a fixed value of V_{BE}, contrary to the predictions of device theory. Increased V_{CB} was expected to decrease the carrier recombination in the QNB and to increase the carrier generation at the base-collector junction. Both of these

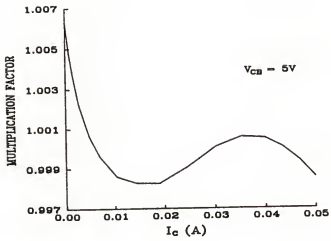


Fig. 4.5 Experimentally extracted multiplication factor versus I_C.

mechanisms would have resulted in decrease of I_B. This anomalous trend, which was more pronounced at higher V_{BE}, was observed not only in IC transistors, but also in discrete transistors. Second, the simulated M is too large in the very high-I_C region. This error can be attributed to an overestimation of α_n in (4.3) at very high E. The error can be removed by using a more physical $\alpha_n(E)$ model, but this may not be worthwhile because even the highest operating currents today are well below the values for which large discrepancies exist.

CHAPTER 5
MMSPICE DEVELOPMENT

5.1 Introduction

In the preceding chapters, the development of the semi-numerical charge-based BJT model has been described. To utilize the model for circuit simulation and for mixed-mode device/circuit simulation, it must be implemented in a circuit simulator like SPICE. In this chapter issues related to the implementation of the model in SPICE2 are discussed. The resulting customized version of SPICE2 was named MMSPICE to indicate its mixed-mode simulation capability, one of the main application areas intended for the semi-numerical charge-based BJT model.

First, the relevance of the semi-numerical mixed-mode simulation approach to TCAD is examined. TCAD [Dut86], which comprises process, device, and circuit simulations, has become an essential part of advanced IC development. An effective TCAD system requires properly linked simulators [Mar88]. For this requirement, there is growing interest in "mixed-mode" device/circuit simulators which provide the simultaneous solutions of the coupled circuit nodal equations and carrier transport equations for the active devices in the

circuit. Such integration eliminates the intermediate step in conventional circuit simulation of compact model parameter extraction, which tends to be empirical and inaccurate. A mixed-mode simulator is capable of predicting circuit performances directly from device structures specified in terms of technological parameters pertaining to doping profile and layout. It thus aids optimal device and circuit design for IC figures of merit, e.g., yield.

Most of the existing mixed-mode simulators [Eng82], [May87], [Rol88] have been created by merging a device simulator and a circuit simulator such that numerical device simulation replaces compact device models in the circuit simulator. While this purely numerical approach gives the most accurate simulations possible for specified devices, the required computation can be prohibitively expensive, especially in statistical simulation applications needed for predicting yields. Furthermore, there is a limit to the accuracy in specifying the device structure due to uncertainties in process simulations or in experimental characterizations. Also, some of the physical models needed in the device simulation cannot be unequivocally defined. These observations suggest that more viable mixed-mode simulations should involve a trade-off between the device modeling rigor and the computation time. An example of such a

trade-off is the recently developed "table method" [Oka88] by which an equivalent circuit (for the bipolar transistor) is characterized by current and capacitance tables derived from numerical dc device simulations.

Another method is presented in this dissertation: a semi-numerical mixed-mode simulation methodology whereby the trade-off is achieved without sacrificing much accuracy. It is called a "semi-numerical" simulation because a set of nonlinear analytic equations derived from regional device analyses, which require numerical iterative solution, replaces the current-continuity equations and Poisson's equation which are directly solved in the numerical simulation. Thus, the resulting semi-numerical simulator has its accuracy and computational burden intermediate between a numerical mixed-mode simulator and a circuit simulator. In fact, its computation time is typically less than an order-of-magnitude longer than that of the circuit simulator, but orders-of-magnitude shorter than a numerical mixed-mode simulator.

The novel method is exemplified by MMSPICE, a semi-numerical mixed-mode simulator that has been developed for advanced bipolar technology. MMSPICE has utility as a circuit simulator as well as a semi-numerical mixed-mode simulator for TCAD applications. The semi-numerical simulator is

created by implementing a physics-based bipolar transistor model in lieu of a built-in compact device model in the circuit simulator SPICE2. However, the difference between this simulator and a conventional circuit simulator is more than an improvement in device model accuracy. The more significant difference is that the former can accept key technological parameters of the device as input whereas the latter generally cannot.

In Section 5.2, the techniques of ensuring the current/charge continuity in integrating different operating modes in the BJT model are discussed. Then, in Section 5.3, the numerical aspects concerning the implicit nature of the model equations are discussed. In Section 5.4, the SPICE2 source code modifications done to produce MMSPICE are discussed. Finally, in Section 5.5, the utility of MMSPICE in TCAD applications is demonstrated, and benchmarks relative to existing device/circuit simulators are presented.

5.2 Model Integration

An important requirement imposed on a device model intended for circuit simulation is that the currents and charges in the model be defined continuously in terms of controlling node voltages. This makes it difficult to implement the charge-based BJT model because the model is

defined through different analytical formulations depending on the operating conditions in the collector. Furthermore, any mathematical manipulations that may be required to ensure the continuity of currents is made difficult by the implicit nature of the model. These problems have tended to hamper implementation into a circuit simulator of physics-based device models derived by the regional analysis approach. In this section, a technique is discussed that facilitates the task of ensuring current/charge continuity while developing and refining the algorithm for such models.

The behavior of I_C is focused on because the continuity of all other currents and charges depends on the continuity of I_C. Since the task of the BJT model routine in a circuit simulator is to return currents and (trans)conductances to the nodal analysis for the two given node voltages V_{BE} and V_{BC}, a so-called voltage-plane technique was used in which the "modal boundaries" between two different operating modes are plotted on the V_{BE}-V_{BC} plane. This technique provides a useful visual aid in merging disparate operating modes into an integrated whole by systematically ensuring the model continuity at the modal boundaries.

The major modal boundaries can be plotted on the voltage-plane as shown in Fig. 5.1 based on the model algorithm in Fig. 3.3 and equations (3.10), (3.13), and

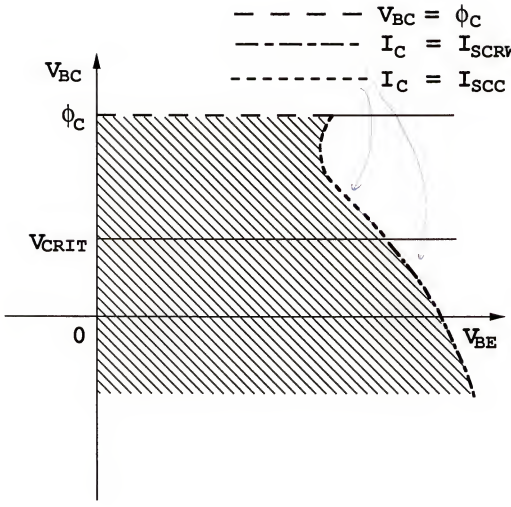


Fig. 5.1 Modal boundaries on the V_{BE}-V_{BC} plane. Dashed lines denote the boundaries between the two distinct operating modes; the cross-hatched area indicates the existence of a space-charge region at the base-collector junction.

(3.15) in Chapter 3. The cross-hatched area in Fig. 5.1 is where a space-charge region exists at the base-collector junction. Thus, potential model discontinuities can result at the modal boundaries denoted by the dashed lines in Fig. 5.1. There are also two minor boundaries that can cause discontinuities. The first of these boundaries is defined by $I_C = I_O$ for $V_{CRIT} < V_{BC} < \phi_C$, because a current-induced space-charge region forms for $I_C > I_O$. The other boundary is constituted by $V_{BC} = \phi_C$, because the entire epitaxial collector is quasi-neutral for $V_{BC} > \phi_C$.

Note that the current, and the regional charges, can be made continuous at the boundaries by ensuring the continuity of n_C and x_C. So, the behaviors of n_C and x_C as defined by corresponding operating modes were carefully checked at all modal boundaries as the operating point approaches the boundaries from either side. For example, a minor discontinuity can result at $I_C = I_{SCRWEPI}$ (see Subsection 3.2.2) if the integral of the electric field in the current-induced space-charge region is equated to the quasi-Fermi potential drop across the region, i.e., $V_{BCI} - V_{BC}$. This discontinuity can be avoided by neglecting the integral of electric field in the quasi-neutral portion of the collector and consistently equating the integral of electric field in the space-charge region to $\phi_C - V_{BC}$. However, a discontinuity

in n_C was inevitable at $I_C = I_{SCC}$ (see Appendix A), requiring a localized empirical smoothing in the region $I_C < I_{SCC}$ to avoid any model discontinuity at the boundary. Detailed description of the methods used to ensure the current/charge continuity at all modal boundaries is provided in Appendix A.

5.3 Numerical Techniques for Model Solution

In this section, the numerical techniques are discussed that are used to solve the model equations. The classical method for numerically obtaining the root of a nonlinear equation is the Newton-Raphson (N-R) iterative method. Even though the N-R method is straightforward and generally has a rapid convergence property, it is not without some drawbacks. First, it requires the derivative of the function. This is disadvantageous when the derivative is difficult to derive analytically or is computationally intensive. Second, it can have convergence difficulties when the function varies abruptly near its root.

To illustrate the numerical problems that are being dealt with, the model is represented, based on (3.18) and Fig. 3.3, as

$$F(I_C) = f(x_E , n_E , x_C , n_C) - I_C = 0. \qquad (5.1)$$

As described in Chapter 3, n_C and x_C are related to I_C in complicated fashions depending on the collector condition.

For example, n_C is defined as a simple function of I_C, assuming saturation of drift velocity, when there exists a space-charge region at the base-collector junction. But when the entire epitaxial collector is quasi-neutral, n_C is defined as a function of the quasi-Fermi potential separation which is in turn implicitly defined by I_C. This implicit nature of the model makes it virtually impossible to directly use the N-R method because of the difficulty in obtaining the required derivative, dF/dI_C.

An alternative to the N-R method is the secant iterative method [Pre86] in which the derivative is approximated by a difference formula in terms of the current and preceding iterates. The secant method, like the N-R method, has occasional difficulties in convergence. There are several alternative methods devised to avoid convergence problems. All of these methods incorporate some means of bracketing which ensures that the iterate remains within some specified interval. One of the commonly used algorithms is the Van Wijngaarden-Dekker-Brent (Brent) method [Pre86] which combines root bracketing, bisection, and inverse quadratic interpolation to converge from the neighborhood of a root. Even though this algorithm does not require the derivative of the function and is guaranteed to converge, it can be computationally inefficient when the function is only rarely

abrupt around its zero crossing. Therefore, the secant method is modified such that bisection is automatically invoked only when a new guess falls outside the specified bracket. In Fig. 5.2, the algorithm for this hybrid secant-bisection method is flowcharted. This algorithm, which defines the iterate control in Fig. 3.3, utilizes the property that $F(I_C)$ in (5.1) is a monotonic function. In simulations where the secant method alone would lead to convergence problems, this hybrid method was found to be effective and required approximately 20% less computation than the Brent method.

5.4 SPICE Code Modification

The BJT model was implemented in SPICE2G.6 via substantive modifications of the source code, creating MMSPICE. The BJT model is a stand-alone, Z-prefix element in MMSPICE, and is accessed similarly to other semiconductor device models written into SPICE2 [Nag75]. The source-code modification comprises two new device routines, QBBJT and QBCT, and changes in 14 of the original routines in the simulator to accept the new model. The BJT model is kept separate from the original (Gummel-Poon) BJT model in SPICE2, and hence the new model is portable to other SPICE-based circuit simulators.

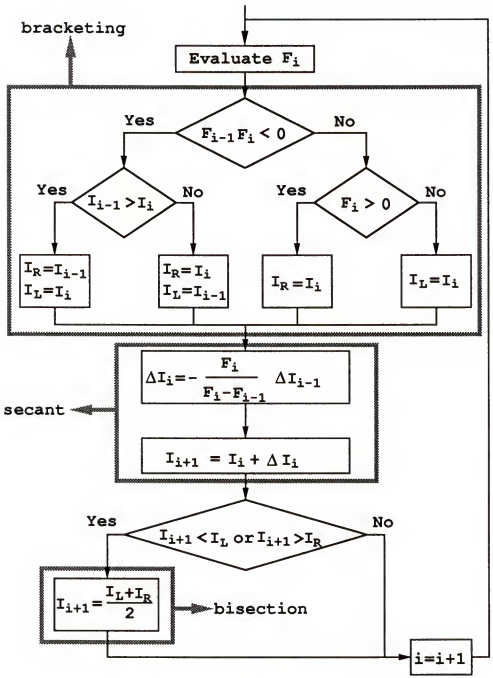


Fig. 5.2 A flowchart of the hybrid secant-bisection method used in the BJT model algorithm in QBCT. The collector current, I_C, is denoted by I in the flowchart.

The two separate device routines are required because of the implicit nature of the BJT model. The numerical solution of the implicit system of nonlinear equations describing the device behavior, flowcharted in Fig. 3.3, is encoded in the QBCT routine. For passed values of V_{BE} and V_{BC}, the routine solves this system using the techniques discussed in Section 5.3. The QBCT routine returns the operating-point currents and regional charges to the QBBJT routine.

Because I_C is defined by an implicit system of equations, analytic (voltage-) derivatives do not exist for the currents and charges through all regions of operation. Thus, numerical (divided-difference) approximations for the derivatives are used in the Newton-Raphson iterations of SPICE2. The numerical derivatives are calculated in the QBBJT routine via successive calls of the QBCT routine with perturbed values of V_{BE} and V_{BC}. The structure of the routine QBBJT is illustrated in Fig. 5.3 The routine initializes variables used in the internal iterations of the QBCT routine based on previous values of the variables and on their dependences of V_{BE} and V_{BC}. Also the QBBJT routine links the model to the nodal analysis by doing other initializations, iterate-limiting, and admittance-matrix load operations, accounting for the extrinsic series resistances shown in Fig. 3.5. The details of the modifications for the other routines

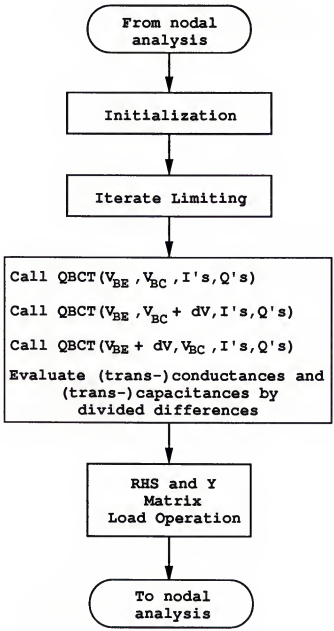

From nodal analysis
Initialization
Iterate Limiting
Call QBCT(V_{BE}, V_{BC}, I's, Q's)
Call QBCT(V_{BE}, V_{BC} + dV, I's, Q's)
Call QBCT(V_{BE} + dV, V_{BC}, I's, Q's)
Evaluate (trans-) conductances and (trans-) capacitances by divided differences
RHS and Y Matrix Load Operation
To nodal analysis

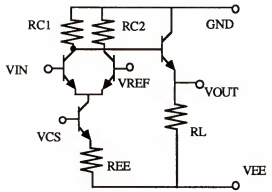
Fig. 5.3 A flowchart of the QBBJT routine. Three calls of the QBCT routine are needed for numerical evaluation of required derivatives.

are described in [Jeo89]. MMSPICE, with the one-dimensional QBBJT/QBCT device routine structure, forms a semi-numerical two- or three-level (depending on the operating mode) Newton(-like) mixed-mode device/circuit simulator.

5.5 Demonstration/Benchmarking

Examples of dc simulation by MMSPICE were presented together with corresponding measurements in Chapter 3. Here, the use of MMSPICE is demonstrated for transient simulations and its potential utility in TCAD applications are indicated. One circuit chosen for simulation is an ECL inverter, which is the basic building block of a ring oscillator commonly used for benchmarking high-speed logic gates. Fig. 5.4 shows the circuit diagram of the inverter and a typical MMSPICE input file for transient simulation of the circuit.

The general input format for using the BJT device model is essentially the same as that for the standard Gummel-Poon-based model in SPICE2. The device card starts with a Z-prefix element name and contains layout-dependent parameters which can differ from device to device on the same IC chip. At the end of the device card, initial conditions for device node voltages can be supplied as an option. These initial conditions can be useful in easing convergence difficulties. The model card contains process-dependent parameters that are




```
THIS IS A CIRCUIT FILE FOR AN ECL INVERTER
Z1 2 1 3 9 NPNMOD AE=3.8P AC=9.5P IC = 0.7 1.0
Z2 5 4 3 9 NPNMOD AE=3.8P AC=9.5P IC = 0.7 1.0
Z3 3 6 7 9 NPNMOD AE=3.8P AC=9.5P IC = 0.7 1.0
Z4 0 2 8 9 NPNMOD AE=3.8P AC=9.5P IC = 0.7 1.0
RC1 0 2 500
RC2 0 5 500
REE 7 9 1840
RL  8 9 4090
VIN  1 0 PULSE(-1.36 -0.86 100P 50P 50P 200P 0.5N)
VREF 4 0 -1.11
VEE  9 0 -5.2
VCS  6 0 -2.5
.MODEL NPNMOD QBNPN (UNEPI=1000
+ NEPI=2.0E16 WEPI=0.375E-6
+ VS=1.0E7 WBM=1.5E-7 JEO=1.0E-8 JSEO=1.46E-4 NEB=2.2
+ TB=1.0E-6 TC=1.0E-5 WSEO=2.0E-8 PE=1.0 ME=0.4 PC=0.7
+ NAO=1.0E18 ETA=4.0 CIF=0.5 CIR=0.5 RC=73 RB=70 RE=10
+ FB=0.5 FC=0.5 TE=1.0E-9 DNB=8.626 VJS=0.8 MJS=0.5
+ CJS=0.0056 CRBI=0 AI=0.0E6 BI=0.0E6)
.TRAN 5P 500P
.PLOT TRAN V(1) V(8) (-1.5,-0.5)
.END
```

Fig. 5.4 Circuit diagram of an ECL inverter and corresponding input file for MMSPICE.

in general common to all the BJT devices on the same chip. Key model parameters are technological and can be evaluated [Gre89] from process information, e.g., SUPREM-3 [Tec86] output. Most of the other parameters are closely related to the device physics and are thus tightly constrained in value based on physical models [Gre89]. These parameters can be fine-tuned to a particular fabrication process through experimental characterization of test devices. When test devices are not available, numerical device simulations can be used in refining the parameters.

Now the use of MMSPICE is exemplified in optimizing device designs for the ECL inverter, based on a representative advanced bipolar technology [Zde87]. The goal in this example is to determine the optimum epitaxial collector thickness that minimizes the propagation delay without incurring breakdown at the operating voltage. To evaluate the propagation delay for unit fan-in and fan-out conditions, a three-stage inverter comprising a cascade of three gates was simulated. Both turn-on delay and turn-off delay were estimated by calculating the time interval between the midpoints of waveform transitions at the input and output of the middle gate, illustrated in Fig. 5.5. Since the turn-on delay is in general different from the turn-off delay, the average of the two delays was taken to be the propagation

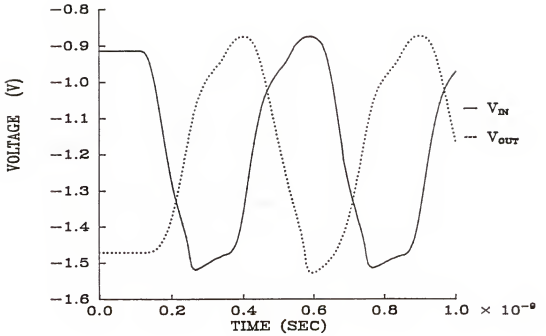


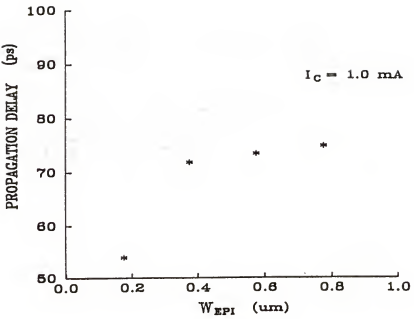
Fig. 5.5 MMSPICE-simulated switching waveforms of an ECL inverter stage

delay. The predicted propagation delay versus epitaxial collector thickness, defined by the complex charge dynamics in the BJT model, is plotted in Fig. 5.6(a). The common-emitter breakdown characteristics, governed by the impact ionization as discussed in Chapter 4, were simulated for a single transistor. The predicted breakdown voltage versus the collector thickness is plotted in Fig. 5.6(b). These results confirm the expected design tradeoff between propagation delay and breakdown characteristics. They correlate well with measurements and demonstrate the TCAD utility of MMSPICE.

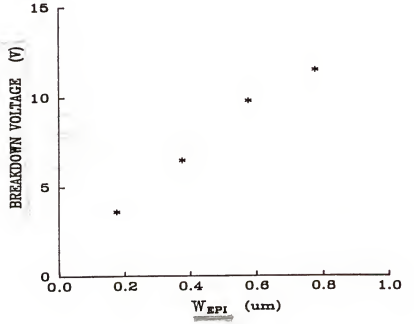
Even though this kind of design optimization can also be achieved by using a purely numerical mixed-mode simulation, a semi-numerical approach is computationally more efficient and thus offers a viable alternative when the simulation time is at a premium. The amount of computation can increase significantly when relatively complicated circuits are simulated or when repetitive simulations are required for sensitivity analyses or statistical parametric study in optimizing yields.

Now the run-time performances of MMSPICE are considered. Benchmarking MMSPICE is difficult due to the lack of any existing simulator of the same kind. So, comparisons made against SPICE2, which is a well-benchmarked program,

Fig. 5.6 (a) MMSPICE-simulated ECL gate propagation delay vs epitaxial collector thickness
(b) MMSPICE-simulated BJT common-emitter breakdown voltage vs epitaxial collector thickness.


PROPAGATION DELAY (ps)
100
90
80
70
60
50
I_C = 1.0 mA
0.0
0.2
0.4
0.6
0.8
1.0
W_{EPI} (um)

(a)


BREAKDOWN VOLTAGE (V)
15
10
5
0
0.0
0.2
0.4
0.6
0.8
1.0
W_{EPI} (um)

(b)

were mainly relied on. Another reference chosen to compare MMSPICE against is CODECS [May87], [May88], which is the first numerical mixed-mode device/circuit simulator available in the public domain. To utilize the benchmark results for CODECS against SPICE3, all the bipolar transistor circuits used in the work [May88] were simulated by MMSPICE on a Sun 3/60 workstation. These circuits include a Colpitts oscillator, a MECL III Motorola ECL gate, an astable multivibrator, a single-stage RTL inverter, a four-stage RTL inverter chain, and a voltage-controlled oscillator. The MMSPICE input files for these benchmark circuits are listed in Appendix C. The model parameters used in the simulations were evaluated based on typical advanced bipolar device structures [Zde87]. The key model parameters are listed in Table 5.1. Table 5.2 lists the run-time performances of MMSPICE relative to SPICE2 and CODECS, based on the assumption that SPICE2 and SPICE3 have comparable run times. Overall, MMSPICE is less than an order-of-magnitude slower than SPICE2, but faster than CODECS by more than an order-of-magnitude. These results demonstrate a significant reduction in computation time when the semi-numerical method is used for device/circuit simulation instead of conventional numerical methods. Since MMSPICE has yet to be optimized in

TABLE 5.1 KEY MODEL PARAMETERS USED FOR BENCHMARKING

Name	Description	Value
NEPI	Epitaxial doping density	2.0×10^{16} cm^{-3}
WEPI	Epitaxial collector thickness	3.75×10^{-7} m
WBM	Metallurgical base width	1.5×10^{-7} m
NAO	Extrapolated peak base doping density	1.0×10^{18} cm^{-3}
ETA	Base doping gradient factor	4.0

TABLE 5.2 MMSPICE/SPICE AND CODECS/MMSPICE RELATIVE RUN-TIMES

Circuits	MMSPICE/SPICE	CODECS/MMSPICE
Colpitts oscillator	3.77	42.44
MECL III Motorola ECL gate	4.98	41.16
Astable multivibrator	8.38	13.6
RTL inverter	5.60	-
Four-stage RTL inverter chain	6.50	38.15
Voltage controlled oscillator	4.11	36.5

TABLE 5.3 BIAS-DEPENDENT AGGREGATE RUN-TIMES FOR MMSPICE

V_{BE} (V)	V_{BC} = 0 V	V_{BC} = -2 V
0.70 - 0.75	1.47	1.25
0.75 - 0.80	1.53	1.48
0.80 - 0.85	1.75	1.82
0.85 - 0.90	2.58	1.73
0.90 - 0.95	7.67	6.93

its numerical solution algorithm, the ultimate benchmarking is expected to be even more favorable.

As shown in Table 5.2, MMSPICE tends to be computationally less efficient when the transistors in the circuit enter into quasi-saturation or saturation regions as in astable multivibrators or RTL circuits. This is to be expected since the solution of the model equations is more complex in these regions of operation. To investigate this problem further, dc simulations were performed as a function of V_{BE} and V_{BC}. Table 5.3 summarizes the accumulated analysis time for 50mV intervals of V_{BE}, which were obtained by performing dc transfer analysis in 1mV steps of V_{BE}. Corresponding SPICE2 simulations show that the CPU time is approximately 0.54 sec for each interval, nearly independent of bias. At high V_{BE} in MMSPICE, the internal forward bias of the base-collector junction causes the base and collector analyses to be linked through the quasi-Fermi potential separation. This requires an iterative solution for n_C as a function of I_C and V_{BC} alluded to in Section 5.4, which increases the simulation time.

Even though a semi-numerical mixed-mode simulation runs faster than its numerical counterpart, it is not as robust against structural changes of the simulated device. Drastic changes in the device structure could require a major

modification of the device models. In many practical applications, the application-specific semi-numerical simulator and its numerical counterpart are expected to play complementary roles. One example is the use of a numerical simulator in evaluating some input parameters of the semi-numerical simulation. By incorporating an input preprocessor that evaluates the model parameters from SUPREM-3 output, MMSPICE can be a basis for a process/device/circuit mixed-mode simulator for general TCAD applications [Gre89].

CHAPTER 6
SUMMARY AND SUGGESTIONS FOR FUTURE WORK

In this dissertation, a new approach of modeling advanced bipolar transistors has been presented which can be used not only for circuit simulation, but also for a semi-numerical device simulation. The charge-based model was derived based on regional analyses involving technological parameters, and can physically account for high-current effects, impact ionization, and nonreciprocal capacitances. With due considerations given to the unique implementation problems, the model was written into the source code of SPICE2 to create MMSPICE. The final version of the code, MMSPICE-1.1 [Jeo89], evolved from this dissertation, is available at the University of Florida. A parameter optimization program was written to expedite model parameter extractions. MMSPICE was benchmarked and the utility of this semi-numerical device/circuit simulator for technology CAD was demonstrated through simulations.

The following tasks are suggested as future tasks to make the simulator more useful.

(1) For the model to be applicable in all regions of transistor operation, the inverse-active mode and hard saturation region must be properly modeled. The key

assumptions used for the base and collector regional analyses in MMSPICE-1.1 become invalid in these operating regimes.

(2) To better approximate the actual doping profile in the base, functional representations more complex than a simple exponential can be used.

(3) To make the model scalable, multidimensional current flows must be properly accounted for.

(4) The transient portion of the model needs to be verified through small-signal s-parameter and large-signal transient measurements.

(5) Numerical aspects associated with unique modeling approach need to be investigated further for potential improvements in convergence and numerical efficiency.

(6) For a fully automated parameter extraction, a driver routine for the measurement equipment, e.g., semiconductor parameter analyzer, must be written.

(7) For effective TCAD, e.g., for applications in design for manufacturability, a process simulator needs to be interfaced to MMSPICE, e.g. as in SUMM [Gre89].

APPENDIX A
DETAILED MODEL ALGORITHM

Two new routines, QBBJT and QBCT, were added to original SPICE2G.6 code to generate MMSPICE. The QBBJT routine closely follows the structure of the original SPICE2 routine BJT, except that all the model calculations are now done in a separate routine QBCT. The model algorithm, which largely dictates the structure of QBCT routine, was discussed in Subsection 3.2.3. Here, the structure of QBCT routine is discussed in more detail to facilitate future source-code modifications needed for debugging and model extensions.

The task of the QBCT routine is to return I_C, I_B, Q_{BE}, and Q_{BC} for the two input node voltages V_{BE} and V_{BC}. The following observations influenced the overall structuring of QBCT. First, the solution of the model algorithm directly provides only I_C. Evaluation of Q_{BE}, Q_{BC}, and I_B, including the impact-ionization-generated current, needs further calculations based on the internal solutions of the model. Thus, a post-processing is needed following the solution of the model algorithm. Since I_B depends on stored charges as well as possible impact ionization, it is calculated last. Second, the collector analysis plays a key role in

determining the major branching of the program flow depending on the value of V_{BC}, because V_{BC}, together with I_C, determines the operating condition of the collector, and hence the pertinent mode of physical linkage to the base transport. The base analysis itself is independent of a particular linkage mode and thus can be shared to avoid duplication of the code. Lastly, all the constants or quantities that can be calculated outside the iterative loop should be done so for numerical efficiency.

The resulting structure of QBCT routine is shown in Fig. A.1. First, model parameters must be initialized based on the value of the pointer LOC passed from the QBBJT routine. Then, constants, critical currents and voltages, and V_{BE}-dependent terms are calculated. Next, the main model solution loop begins. An initial guess is made for I_C^C, the collector transport current. This I_C^C and the given node voltage V_{BC} completely define the device operating modes as distinguished by their different collector charge conditions. Boundaries between these operating modes are defined differently depending on the value of V_{BC} with respect to ϕ_C and V_{CRIT}. Thus, three major branches result in the program structure as shown in Fig. A.1. Additional minor branches also need to be

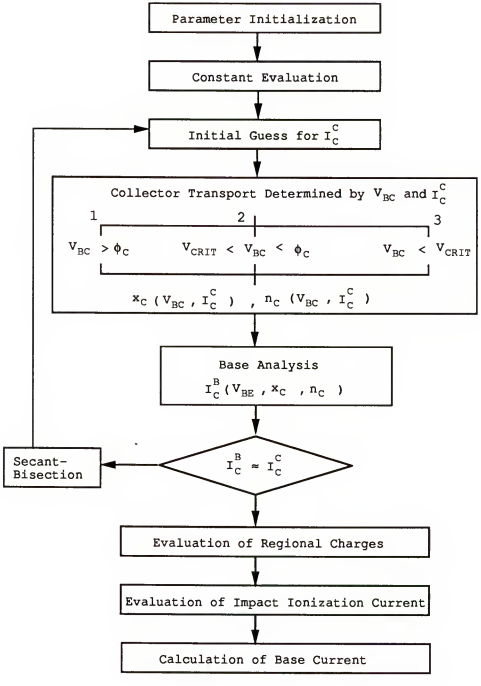


Fig. A.1 Detailed flowchart for QBCT routine.

defined as shown in Tables A.1 and A.2, and as discussed below.

In Branch 1 ($V_{BC} > \phi_C$), the entire epitaxial collector is quasi-neutral and no further branching needs to be defined. In Branch 2 ($V_{CRIT} < V_{BC} < \phi_C$), W_{SCC} contracts with I_C as indicated by (3.8). For this branch, (3.9) defines the onset condition for the ohmic quasi-saturation. Furthermore another operating mode is defined by the onset of nonohmic quasi-saturation when I_C exceeds I_O. Thus, three minor branches are defined for Branch 2 as described in Table A.1: $I_C < I_{SCC}$ (Branch 2a), $I_{SCC} < I_C < I_O$ (Branch 2b), and $I_C > I_O$ (Branch 2c). In Branch 3 ($V_{BC} < V_{CRIT}$), W_{SCC} expands with I_C. For this branch (3.12) defines the value of I_C at which the entire epitaxial collector becomes space-charged. Further increase of I_C causes the onset of non-ohmic quasi-saturation for $I_C > I_{SCRWEPI}$, given by (3.14). Thus, three minor branches result for Branch 3 as described in Table A.2: $I_C < I_{SCCWEPI}$ (Branch 3a), $I_{SCCWEPI} < I_C < I_{SCRWEPI}$ (Branch 3b), and $I_C > I_{SCRWEPI}$ (Branch 3c).

The solution of the collector transport provides n_C and x_C in terms of V_{BC} and I_C^C, as shown in Tables 5.1 and 5.2, based on the collector analysis discussed in Chapter 3. Combining the equations (3.16), (3.17), and (3.26) yields Q_{JC}.

TABLE A.1 SUB-BRANCHES OF BRANCH 2 IN FIG. A.1

Branches	Branch 2-a	Branch 2-b	Branch 2-c
Range of I_C	$I_C < I_{SCC}$	$I_{SCC} < I_C < I_O$	$I_C > I_O$
Charge Condition	Fig. 3.2(a)	Fig. 3.2(b)	Fig. 3.2(c)
n_C	$n_C = I_C/qAv_S$	$n_C(V_{BCO})$	$n_C(V_{BCO})$
x_C	$x_C(Q_{JC})$	$x_C = x_{JC}$	$x_C = x_{JC}$

TABLE A.2 SUB-BRANCHES OF BRANCH 3 IN FIG. A.1

Branches	Branch 3-a	Branch 3-b	Branch 3-c
Range of I_C	$I_C < I_{SCCWEPI}$	$I_C > I_{SCCWEPI}$ $I_C < I_{SCRWEPI}$	$I_C > I_{SCRWEPI}$
Charge Condition	Fig. 3.2(a)	Fig. 3.2(d)	Fig. 3.2(c)
n_C	$n_C = I_C/qAv_S$	$n_C = I_C/qAv_S$	$n_C(V_{BCO})$
x_C	$x_C(Q_{JC})$	$x_C(Q_{JC})$	$x_C = x_{JC}$

Then, (3.25) can be used to derive x_C in terms of Q_{JC}. For Branches 2b, 2c, and 3c, n_C is represented in terms of the quasi-Fermi potential separation V_{BCO}, which can be determined by solving

$$I_C^C - \frac{K_1(V_{BCO}) - K_1(V_{BC}) - \ln\left[\frac{1 + K_1(V_{BCO})}{1 + K_1(V_{BC})}\right] + \frac{V_{BCO}-V_{BC}}{V_T}}{R_{EPI}/V_T} = 0 \quad \text{(A.1)}$$

when the entire epitaxial collector is quasi-neutral, or by solving

$$I_C^C - \frac{K_1(V_{BCO}) - K_1(V_{BCI}) - \ln\left[\frac{1 + K_1(V_{BCO})}{1 + K_1(V_{BCI})}\right] + \frac{V_{BCO}-V_{BCI}}{V_T}}{R_{EPI}/V_T} = 0 \quad \text{(A.2)}$$

when a current-induced SCR exists. The solution of these nonlinear equations requires an additional iteration inside the main loop. This tends to increase the simulation time when the transistor is in saturation or quasi-saturation.

With x_C and n_C available, it is straightforward to next evaluate I_C^B by (3.5). If I_C^B is not close to I_C^C within a specified tolerance, the new guess for I_C^C is estimated using the secant-bisection method described in Section 5.3. Otherwise, the main model solution loop is exited. Then, regional charges, impact-ionization-generated current, and base current are evaluated in sequence.

The model incorporates different operating modes as discussed. Consequently, one of the practical issues in implementing such a model for computer simulation is the seamless integration of constituent operating modes. This involved, for some types of discontinuities, systematic removals of such continuities based on the understanding of model behavior at the modal boundaries. However, some empirical smoothing was inevitable for narrow regions near some of the modal boundaries. Initially, attempts were made to eliminate any discontinuities by tracing them to the basic model equations and by redefining pertinent model equations. When this was impossible, a localized smoothing was imposed on the pertinent variables on one side of the boundary.

First, the systematic removal of model discontinuities is discussed. One such removal was briefly exemplified in Subsection 5.2.2. Another discontinuity would result in x_C at the boundary defined by $I_C = I_{SCRWEPI}$ in Branch 2 due to the abrupt transition in operating mode from (c) to (d) in Fig. 3.2. This comes from, in essence, the abrupt increase of electric field at the junction from 0 to E_S due to simplifying assumptions used in the model. This problem can be resolved by neglecting the junction electric field below E_S, or the corresponding Q_{JC}, for calculating x_C.

Another discontinuity arises in connection with assuming the abrupt transformation of the unmodulated epitaxial collector region into a current-induced SCR when I_C exceeds I_O. This corresponds to the transition from (b) to (c) in Fig. 3.2. To resolve this discontinuity, (2.1) is applied only to the conductivity-modulated region as is done to derive (2.9). Thus, (2.1) is transformed into

$$I_C = \frac{K_1(V_{BCO}) - K_1(V_{BCI}) - \ln\left[\frac{1 + K_1(V_{BCO})}{1 + K_1(V_{BCI})}\right] + \frac{V_{BCO}-V_{BCI}}{V_T}}{R'_{EPI}/V_T} \qquad (A.3)$$

where $$R'_{EPI} = R_{EPI}\frac{W_{EPI} - W_{UNMOD}}{W_{EPI}} \qquad (A.4)$$

and $$W_{UNMOD} = \frac{(\phi_C - V_{BC})W_{EPI}}{R_{EPI}I_C}. \qquad (A.5)$$

At $I_C = I_O$, (A.5) reduces to

$$W_{UNMOD} = \frac{\phi_C - V_{BC}}{E_S}.$$

Thus, (A.3) and (2.9) become equivalent to each other at $I_C = I_O$ if V_{BCI} is defined identically. However, V_{BCI} is not well-defined at $I_C = I_O$ as can be seen in (2.8). This means the model is not sensitive to the value of V_{BCI} at $I_C = I_O$. Here, V_{BCI} was arbitrarily forced to be zero at $I_C = I_O$.

Now the model continuity at $V_{BC} = \phi_C$ can be examined by rewriting (A.3). Since R'_{EPI} is a function of I_C, (A.3) can be solved for I_C:

$$I_C = \frac{K_1(V_{BCO}) - K_1(V_{BCI}) - \ln\left[\frac{1+K_1(V_{BCO})}{1+K_1(V_{BCI})}\right] + \frac{V_{BCO}-V_{BCI}+\phi_C-V_{BC}}{V_T}}{R_{EPI}/V_T} \quad . \quad (A.6)$$

Note that the model continuity can be ensured at the boundary defined by $V_{BC} = \phi_C$ if V_{BCI} is equated to V_{BC} at points away from the boundary.

The continuity of electron density at the $I_C = I_{SCC}$ boundary could not be removed by manipulating the model equations. Thus, an exponential decaying function was superimposed on the expression for n_C in the region where $I_C < I_{SCC}$. Similar smoothing was done to ensure the continuity at $V_{BC} = \phi_C$ for $I_C > I_O$.

APPENDIX B
PARAMETER OPTIMIZATION/EXTRACTION

Simulation accuracy is critically dependent on the proper choice of model parameters. So, effective use of the simulator demands a well-developed method of parameter extraction. There are, in general, two alternative approaches to measurement-based parameter extraction. In the sequential method, parameters are isolated into small groups based on physical origins of the parameters, and each group of parameters is extracted sequentially assuming the rest of the parameters are accurate. This method has the advantage of facilitating the retention of the physical nature of the parameters and of being numerically less intensive compared to the global optimization method. In the global method, a general purpose optimization technique is used to obtain a complete set of model parameters in a single operation. This is accomplished by simultaneously adjusting all the parameters to produce a nonlinear least-square fit of the model to the measured data. Even though this method is more time-consuming, it has the advantage of accounting for the interactions among parameters.

The parameter-extraction method described in Chapter 3 falls into the first category and was designed to be usable without the aid of an optimizing program. The same algorithm can also be directly applied to the extraction with the optimization program. However, the extraction procedures can be simplified by making use of the powerful optimization program. There are many alternative algorithms to be explored, but no exhaustive attempt has been made to sift through all the possibilities. Here one exemplary algorithm is outlined.

The sequential method was preferred in consideration of numerical intensiveness and possible convergence difficulties associated with the global method, as well as the physical nature of the model. In the sequential method, optimization must start with the most influential group of parameters and proceed in decreasing order of their influences on the device characteristics. Since physical parameters have predominant effects on the device characteristics, they must be optimized first. The next important group of parameters is the semi-empirical parameters. These parameters can be further subgrouped into β-related and non-β-related parameters. The former can be first optimized based on $\beta(V_{BE},V_{BC})$ characteristics and then the latter group is optimized based on the combined data set of $I_C(I_B,V_{CE})$ and $\beta(V_{BE},V_{BC})$ data.

This algorithm is restricted to extraction of physical and semi-physical parameters based on dc measurements. Extraction of empirical parameters will require an ac data base to drive the optimizer.

To optimize the parameter-extraction algorithm, the physical nature of the model must be further exploited. Based on the physical understanding, certain parameters can be predicted to exert pronounced influences in particular regions of operation. For example, simulation will be insensitive to A_C and τ_C except in quasi-saturation or saturation. Another consideration related to the physical nature of the model is that reasonable constraints can be placed on inter-relationships among certain parameters and on parameter values. For example, base doping density at the base-collector junction cannot be smaller than the epitaxial doping density.

Now the mathematical basis underlying the parameter optimization is briefly discussed. Suppose that N data points (x_i, y_i), $i=1,\ldots,N$, are fitted to a model which has a nonlinear dependence on the set of M adjustable parameters p_k, $k=1,\ldots,M$. The model implicitly defines the functional relationship between the dependent and independent variables,

$$y(\underline{x},\underline{p}) = y(\underline{x};p_1,\ldots,p_M) \ . \qquad (B.1)$$

The merit function chosen is chi-square or weighted least-square given by

$$\chi^2(\underline{p}) \equiv \sum_{i=1}^{N} \left[\frac{y_i - y(x_i,\underline{p})}{\sigma_i}\right]^2 \qquad (B.2)$$

where σ_i is the standard deviation for each data point. Even when the standard deviation is not known, σ_i can be assumed to weight a particular portion of data or can be set to y_i to estimate relative error instead of absolute error.

The optimization problem is to minimize $\chi^2(\underline{p})$ by adjusting $\underline{p}$. When the current $\chi^2(\underline{p}^j)$ is sufficiently close to the minimum, $\chi^2(\underline{p})$ can be well approximated by a quadratic form:

$$\chi^2(\underline{p}^j + \Delta\underline{p}) \approx \chi^2(\underline{p}^j) + \Delta\underline{p}^T \underline{g}^j + \frac{1}{2} \Delta\underline{p}^T \underline{G}^j \Delta\underline{p} \qquad (B.3)$$

where $\underline{g}^j \equiv \nabla\chi^2(\underline{p}^j)$ and $\underline{G}^j \equiv \nabla^2\chi^2(\underline{p}^j)$.

At the extremum, the gradient of χ^2 will vanish. From (B.3), the gradient of χ^2 with respect to $\Delta\underline{p}$ can easily be calculated as

$$\nabla\chi^2|_{\Delta\underline{p}} = \underline{g}^j + \underline{G}^j \Delta\underline{p}. \qquad (B.4)$$

Setting (B.4) to zero yields

$$\underline{p}^{j+1} = \underline{p}^{j} - [G^{j}]^{-1} g^{j}. \tag{B.5}$$

Now let us write out g and G defined in (B.3):

$$g_k \equiv \frac{\partial \chi^2}{\partial p_k} = -2 \sum_{i=1}^{N} \frac{y_i - y(x_i;\underline{p})}{\sigma_i^2} \frac{\partial y(x_i;\underline{p})}{\partial p_k} \tag{B.6}$$

$$G_{kl} \equiv \frac{\partial^2 \chi^2}{\partial p_k \partial p_l}$$

$$= 2 \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \{ \frac{\partial y(x_i;\underline{p})}{\partial p_k} \frac{\partial y(x_i;\underline{p})}{\partial p_l} - [y_i - y(x_i;\underline{p})] \frac{\partial^2 y_i(x_i;\underline{p})}{\partial p_l \partial p_k} \}. \tag{B.7}$$

In practice, the second derivative in (B.7) is usually small, when summed over i, and neglected. This is the Gauss-Newton method, which gives rapid quadratic convergence near the minimum. But far away from the minimum, (B.3) and (B.4) become a poor approximation and the method is prone to convergence problems. At points far away from the minimum, the best alternative is just to take a step down the gradient, as in the steepest-descent method. In other words,

$$\underline{p}^{j+1} = \underline{p}^{j} - \text{constant} \times \underline{g}^{j} \tag{B.8}$$

where the constant is small enough not to exhaust the downhill direction.

By clever choice of the constant in (B.8), the Levenberg-Marquardt (L-M) [Bro72] method combines (B.5) and (B.8) to provide a continuous switching from the steepest descent at points far away from the minimum to the Gauss-Newton method as the minimum is approached. Thus, in the L-M method, (B.8) is rewritten as

$$\Delta \underline{p}_k = \frac{1}{\lambda \frac{\partial^2 \chi^2}{\partial \underline{p}_k^2}} \frac{\partial \chi^2}{\partial \underline{p}_k} \quad . \tag{B.9}$$

Then, (B.9) can be merged with (B.5) by redefining the diagonal elements of G by

$$\frac{\partial^2 \chi^2}{\partial \underline{p}_k^2} \rightarrow (1 + \lambda) \frac{\partial^2 \chi^2}{\partial \underline{p}_k^2} \quad . \tag{B.10}$$

When λ is very large, the L-M method reduces to the steepest descent method. On the other hand, as λ approaches zero, the L-M method becomes identical to the Gauss-Newton method.

Given an initial guess for the set of fitted parameters p, the method proceeds as follows:

(1) Compute $\chi^2(\underline{p})$ using (B.2).

(2) Pick a modest value for λ, say $\lambda = 0.001$.

(3) Evaluate $\Delta\underline{p}$ using (B.9) modified as indicated in (B.10), and $\chi^2(\underline{p}+\Delta\underline{p})$.

(4) If $\chi^2(\underline{p}+\Delta\underline{p}) \geq \chi^2(\underline{p})$, increase λ by a factor of 10 and go back to (3).

(5) If $\chi^2(\underline{p}+\Delta\underline{p}) < \chi^2(\underline{p})$, decrease λ by a factor of 10 and go back to (3).

The parameter optimization program was implemented as interactive software consisting of four modules as shown in Fig. B.1. The main and L-M modules were written in Fortran, whereas the interface module was written as a shell script. The main module generates prompts for user inputs as to the names of data files, circuit files, and names of parameters to be optimized. It also allows the user to specify initial parameter values. Program flow is transferred from the L-M module to the interface module by using the "system" feature (available in Fortran runtime library) which allows the execution of Unix commands inside a Fortran program. The interface module first sets up a circuit file by properly adjusting parameter values and runs MMSPICE. Then output data are extracted and the program flow is transferred back to the L-M module. Data transfers between the L-M module and the interface module are achieved through file read/write. In the circuit file generated by the interface module, the ALTER

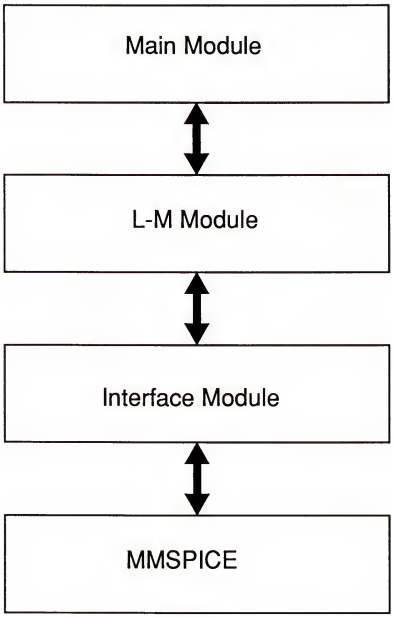


Fig. B.1 Structural chart for parameter optimization program.

statement is used repeatedly for each additional set of parameters to save run-times.

The only significant modifications done to the conventional L-M method were the addition of parameter limiting, based on the physical nature of the model, and the replacement of analytic derivatives by numerical difference approximations, derived from successive simulations.

APPENDIX C
BENCHMARK CIRCUITS

In this appendix, input circuit files used for benchmarking MMSPICE are listed. These circuits are a subset of benchmark circuits used for CODECS [May88]. Circuits containing MOSFETs were excluded. The purpose of using the same benchmark circuits was to enable an indirect benchmarking of MMSPICE against CODECS, as well as SPICE, by comparing performances relative to SPICE and assuming SPICE2 and SPICE3 (used in [May88]) have comparable numerical performances.

```
RTL INVERTER CIRCUIT
VIN 1 0 DC 1 PWL 0 4 1N 0
VCC 12 0 DC 5.0
RC1 12 3 2.5K
RB1 1 2 8K
Z1 3 2 0 0 NPNMOD AE=10P AC=10P WE=10U IC =0.0 4.5
.MODEL NPNMOD QBNPN (UNEPI=1000 NEPI=2.0E16 WEPI=0.375E-6
+ VS=1.0E7 WBM=1.5E-7 JEO=1.0E-8 JSEO=1.46E-4 NEB=2.2
+ TB=1.0E-6 TC=1.0E-5 WSEO=2.0E-8 PE=1.0 ME=0.4 PC=0.7
+ NAO=1.0E18 ETA=4.0 CIF=0.5 CIR=0.2 RC=50 RB=100 RE=5
+ FB=0.5 FC=0.5 TE=0.5E-9 DNB=8.626 PS=0.80 MS=0.50 CJS=0.00
+ CRBI=0.0 AI=0.0E6 BI=0.00E6)
.OPTION ACCT
.TRAN 0.5N 5N
```

```
.PRINT TRAN V(3)
.END

RTL INVERTER CHAIN
VIN 1 0 DC 1 PWL 0 4 1N 5
VCC 12 0 DC 5.0
RC1 12 3 2.5K
RB1 1 2 8K
Z1 3 2 0 0 NPNMOD AE=10P AC=10P WE=10U IC =0.0 4.5
RB2 3 4 8K
RC2 12 5 2.5K
Z2 5 4 0 0 NPNMOD AE=10P AC=10P WE=10U IC =0.8 0.0
RB3 5 6 8K
RC3 12 7 2.5K
Z3 7 6 0 0 NPNMOD AE=10P AC=10P WE=10U IC =0.0 4.5
RB4 7 8 8K
RC4 12 9 2.5K
Z4 9 8 0 0 NPNMOD AE=10P AC=10P WE=10U IC =0.8 0.0
.MODEL NPNMOD QBNPN (UNEPI=1000 NEPI=2.0E16 WEPI=0.375E-6
+ VS=1.0E7 WBM=1.5E-7 JEO=1.0E-8 JSEO=1.46E-4 NEB=2.2
+ TB=1.0E-6 TC=1.0E-5 WSEO=2.0E-8 PE=1.0 ME=0.4 PC=0.7
+ NAO=1.0E18 ETA=4.0 CIF=0.5 CIR=0.2 RC=50 RB=100 RE=5
+ FB=0.5 FC=0.5 TE=0.5E-9 DNB=8.626 PS=0.80 MS=0.50 CJS=0.00
+ CRBI=0.0 AI=0.0E6 BI=0.00E6)
.TRAN 1N 10N
.OPTIONS ACCT
.PRINT TRAN V(3) V(5) V(9)
.END

MECL III CKT
.OPTIONS ACCT
.TRAN 0.2NS 20NS
```

```
VEE 22 0 -6.0
VIN 1 0 PULSE -0.8 -1.8 0.2NS 0.2NS 0.2NS 10NS 20NS
RS 1 2 50
Z1 4 2 6 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
Z2 4 3 6 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
Z3 5 7 6 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
Z4 0 8 7 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
D1 8 9 DMOD
D2 9 10 DMOD
RP1 3 22 50K
RC1 0 4 100
RC2 0 5 112
RE 6 22 380
R1 7 22 2K
R2 0 8 350
R3 10 22 1958
Z5 0 5 11 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
Z6 0 4 12 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
RP2 11 22 560
RP3 12 22 560
Z7 13 12 15 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
Z8 14 16 15 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
RE2 15 22 380
RC3 0 13 100
RC4 0 14 112
Z9 0 17 16 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
R4 16 22 2K
R5 0 17 350
D3 17 18 DMOD
D4 18 19 DMOD
R6 19 22 1958
Z10 0 14 20 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
```

```
Z11 0 13 21 22 NPNMOD AE=100P AC=100P WE=10U IC =0.8 1.0
RP4 20 22 560
RP5 21 22 560
.MODEL DMOD D RS=40 TT=0.1NS CJO=0.9PF N=1 IS=1E-14
+ EG=1.11 VJ=0.8 M=0.5
.MODEL NPNMOD QBNPN (UNEPI=1000 NEPI=2.0E16 WEPI=0.375E-6
+ VS=1.0E7 WBM=1.5E-7 JEO=1.0E-8 JSEO=1.46E-4 NEB=2.2
+ TB=1.0E-6 TC=1.0E-5 WSEO=2.0E-8 PE=1.0 ME=0.4 PC=0.7
+ NAO=1.0E18 ETA=4.0 CIF=0.5 CIR=0.2 RC=20 RB=50 RE=5
+ FB=0.5 FC=0.5 TE=0.5E-9 DNB=8.626 PS=0.80 MS=0.50 CJS=0.00
+ CRBI=0.0 AI=0.0E6 BI=0.00E6)
.PRINT TRAN V(12) V(21)
.END

ASTABLE MULTIVIBRATOR
VIN 5 0 DC 0 PULSE (0 5 0 1U 1U 100U 100U)
VCC 6 0 5.0
RC1 6 1 600
RC2 6 2 600
RB1 6 3 30K
RB2 5 4 30K
C1 1 4 150P
C2 2 3 150P
Z1 1 3 0 0 NPNMOD AE=10P AC=10P WE=10U IC =0.7 1.0
Z2 2 4 0 0 NPNMOD AE=10P AC=10P WE=10U IC =0.0 5.0
.MODEL NPNMOD QBNPN (UNEPI=1000 NEPI=2.0E16 WEPI=0.375E-6
+ VS=1.0E7 WBM=1.5E-7 JEO=1.0E-8 JSEO=1.46E-4 NEB=2.2
+ TB=1.0E-6 TC=1.0E-5 WSEO=2.0E-8 PE=1.0 ME=0.4 PC=0.7
+ NAO=1.0E18 ETA=4.0 CIF=0.5 CIR=0.2 RC=50 RB=100 RE=5
+ FB=0.5 FC=0.5 TE=0.5E-9 DNB=8.626 PS=0.80 MS=0.50 CJS=0.00
+ CRBI=0.0 AI=0.0E6 BI=0.00E6)
.OPTION ACCT
```

```
.TRAN 0.05U 8U 0 0.05U
.PRINT TRAN V(1) V(2) V(3) V(4)
.END

COLPITT'S OSCILLATOR CIRCUIT
.WIDTH IN=80 OUT=80
R1 1 0 1
Z1 2 1 3 0 NPNMOD AE=30P AC=30P WE=10U IC =0.7 1.0
VCC 4 0 5
RL 4 2 750
C1 2 3 500P
C2 4 3 4500P
L1 4 2 5U
RE 3 6 4.65K
VEE 6 0 DC -15 PWL 0 -15 1N -10
.MODEL NPNMOD QBNPN (UNEPI=1000 NEPI=2.0E16 WEPI=0.375E-6
+ VS=1.0E7 WBM=1.5E-7 JEO=1.0E-8 JSEO=1.46E-4 NEB=2.2
+ TB=1.0E-6 TC=1.0E-5 WSEO=2.0E-8 PE=1.0 ME=0.4 PC=0.7
+ NAO=1.0E18 ETA=4.0 CIF=0.5 CIR=0.2 RC=50 RB=100 RE=5
+ FB=0.5 FC=0.5 TE=0.5E-9 DNB=8.626 PS=0.80 MS=0.50 CJS=0.00
+ CRBI=0.0 AI=0.0E6 BI=0.00E6)
.TRAN 30N 12U
.PRINT TRAN V(2)
.OPTION ACCT LIMPTS=1000
.PLOT TRAN V(2)
.END


VOLTAGE CONTROLLED OSCILLATOR
RC1 7 5 1.5K
RC2 7 6 1.5K
Z5 7 7 5 0 NPNMOD AE=30P AC=30P WE=10U IC =0.8 0.8
```

```
Z6 7 7 6 0 NPNMOD AE=30P AC=30P WE=10U IC =0.8 0.8
Z3 7 5 2 0 NPNMOD AE=30P AC=30P WE=10U IC =0.8 1.6
Z4 7 6 1 0 NPNMOD AE=30P AC=30P WE=10U IC =0.8 1.6
IB1 2 0 0.5M
IB2 1 0 0.5M
CB1 2 0 1P
CB2 1 0 1P
Z1 5 1 3 0 NPNMOD AE=30P AC=30P WE=10U IC =0.8 2.0
Z2 6 2 4 0 NPNMOD AE=30P AC=30P WE=10U IC =0.8 2.0
C1 3 4 0.1U
IS1 3 0 DC 2.5M PULSE 2.5M 0.5M 0 1U 1U 50M
IS2 4 0 1M
VCC 7 0 10
.MODEL NPNMOD QBNPN (UNEPI=1000 NEPI=2.0E16 WEPI=0.375E-6
+ VS=1.0E7 WBM=1.5E-7 JEO=1.0E-8 JSEO=1.46E-4 NEB=2.2
+ TB=1.0E-6 TC=1.0E-5 WSEO=2.0E-8 PE=1.0 ME=0.4 PC=0.7
+ NAO=1.0E18 ETA=4.0 CIF=0.5 CIR=0.2 RC=50 RB=100 RE=5
+ FB=0.5 FC=0.5 TE=0.5E-9 DNB=8.626 PS=0.80 MS=0.50 CJS=0.00
+ CRBI=0.0 AI=0.0E6 BI=0.00E6)
.OPTION ACCT
.TRAN 3U 300U 0 3U
.PRINT TRAN V(4)
.END
```

REFERENCES

Bow73 D. L. Bowler and F. A. Lindholm, "High current regimes in transistor collector region," *IEEE Trans. Electron Devices*, vol. ED-20, pp.257-263, Mar. 1973.

Bro72 K. M. Brown and J. E. Dennis, Jr., "Derivative free analogues of Levenberg-Marquardt and Gauss algorithms for nonlinear least-squares approximation," *Numer. Math.*, vol. 18, pp. 289-297, 1972.

Dut86 R. W. Dutton and M. R. Pinto, "The use of computer aids in IC technology evolution," *Proc. IEEE*, vol. 74, pp. 1730-1740, Dec. 1986.

Eng82 W. L. Engl, R. Laur, and H. K. Dirks, "MEDUSA: A simulator for modular circuits," *IEEE Trans. Computer-Aided Design of ICAS*, vol. CAD-1, pp. 85-93, Apr. 1982.

Fos86 J. G. Fossum and S. Veeraraghavan, "Partitioned-charge-based modeling of bipolar transistors for non-quasi-static circuit simulation," *IEEE Electron Device Lett.*, vol. EDL-7, pp. 652-654, Dec. 1986.

Fos87 J. G. Fossum, H. Jeong, D. E. Burk, and M. S. Jo, "A charge-based large-signal bipolar transistor model," presented at the IEEE Bipolar Circuits Technol. Meeting, Minneapolis, MN, Sept. 21-22, 1987.

Fra85 A. F. Franz and G. A. Franz, *BAMBI 1.0 Manual*, Technische Universitat Wien, Austria, 1985.

Get78 I. E. Getreu, *Modeling the Bipolar Transistor*. New York: Elsevier, 1978.

Gra85 H. C. de Graaff and W. J. Kloosterman, "New formulation of the current and charge relations in bipolar transistor modeling for CACD purposes," *IEEE Trans. Electron Devices*, vol. ED-32, pp. 2415-2419, Nov. 1985.

Gre89 K. R. Green, "SUMM", unpublished work, University of Florida, Gainesville, 1989.

Hwa87 B.-Y. Hwang, P. J. Zdebel, R. J. Balda, G. G. Sweeney, and V. de la Torre, "Lateral scaling effects of double-

polysilicon advanced self-aligned bipolar npn transistors," presented at the IEEE Bipolar Circuits Technol. Meeting, Minneapolis, MN, Sept. 21-22, 1987.

Jeo89 H. Jeong, D. K. FitzPatrick, and J. G. Fossum, "MMSPICE-1.1 Reference Manual," University of Florida, Gainesville, FL, Aug. 1989.

Kon86 S. Konaka, Y. Yamamoto, and T. Sakai, "A 30-ps Si bipolar IC using super self-aligned process technology," *IEEE Trans. Electron Devices*, vol. ED-33, pp. 526-531, Apr. 1986.

Kul85 G. M. Kull, L. W. Nagel, S. W. Lee, P. Lloyd, E. J. Prendergast, and H. Dirks, "A unified circuit model for bipolar transistors including quasi-saturation effects," *IEEE Trans. Electron Devices*, vol. ED-32, pp. 1103-1113, June 1985.

Kum75 R. Kumar and L. P. Hunter, "Collector capacitance and high-level injection effects in bipolar transistors," *IEEE Trans. Electron Devices*, vol. 24, pp. 53-69, 1969.

Lio87 J. J. Liou, F. A. Lindholm, and J. S. Park, "Forward-bias capacitance and thickness of p-n junction space-charge region," *IEEE Trans. Electron Devices*, vol. ED-34, pp. 1571-1579, July 1987.

Mar88 J. Mar, "CAD for manufacturing: Why its time is now", *Semiconductor Research Corporation TECHON'88 Conference*, Dallas, Texas, Oct. 12-14, 1988.

May87 K. Mayaram and D. O. Pederson, "CODECS: An object-oriented mixed-level circuit and device simulator", *Proc. IEEE Int. Symp. on Circuits and Systems*, pp. 604-607, May 1987.

May88 K. Mayaram, "CODECS: A mixed-mode circuit and device simulator", Electron Res. Lab. Memo. ERL-M88/71, University of California, Berkeley, Nov. 1988.

Mol56 J. L. Moll and I. M. Ross, "The dependence of transistor parameters on the distribution of base layer resistivity," *Proc. IRE*, vol. 44, pp. 72-78, Jan. 1956.

Nag75 L. W. Nagel, "SPICE2: A computer program to simulate semiconductor circuits", Electron Res. Lab. Memo. ERL-M520, University of California, Berkeley, May 1975.

Oka88 H. Oka, S. Satoh, and N. Nakayama, "Bipolar circuit simulation system using a two-dimensional device simulator", *Simulation of Semiconductor Devices and Processes*, vol. 3, ed. by G. Baccarani and M. Rudan, pp. 127-138, Bologna, Italy, Sept. 26-28, 1988.

Pal69 J. A. Pals and H. C. de Graaff, "On the behavior of the base-collector junction of a transistor at high collector current densities," *Philips Res. Rep.*, vol. 24, pp. 53-69, 1969.

Pre86 W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge, Cambridge University Press, 1986.

Rey75 G. Rey, F. Dupuy, and J. P. Bailbe, "A unified approach to the base widening mechanisms in bipolar transistors," *Solid-State Electron.*, vol. 18, pp. 863-866, 1975.

Rol88 J. G. Rollins and J. Choma, "Mixed-mode PISCES-SPICE coupled circuit and device solver", *IEEE Trans. Computer-Aided Design*, vol. 7. pp 862-867, Aug. 1988.

Sak89 K. Sakui, T. Hasegawa, T. Fuse, S. Watanabe, K. Ohuchi, and F. Masuoka, "A new static memory cell based on the reverse base current effect of bipolar transistors", *IEEE Trans. Electron Devices*, vol. ED-36, pp. 1215-1217, June 1989.

Sho61 W. Shockley, "Problems related to p-n junctions in silicon", *Solid-State Electron.*, vol. 2, pp. 35-67, 1961.

Tas84 T. Tashiro, H. Takemura, T. Kamiya, and F. Tokuyoshi, "An 80 ps ECL circuit with high current density transistor," in *IEDM Tech. Dig.*, pp. 686-689, 1984.

Tec86 Technology Modeling Associates, Inc.,"TMA SUPREM-3 : One-Dimensional Process Analysis Program Version 8636", Palo Alto, CA, Oct. 15, 1986.

Ver87 D. P. Verret and J. E. Brighton, "Two-dimensional effects in the bipolar polysilicon self-aligned transistor," *IEEE Trans. Electron Devices*, vol. ED-34, pp. 2297-2303, Nov. 1987.

Whi69 R. J. Whittier and D. A. Tremere, "Current gain and cutoff frequency falloff at high currents," *IEEE Trans. Electron Devices*, vol. ED-16, pp. 39-57, Jan. 1969.

Yun88 S.-Y. Yung and D. E. Burk, "Informed device design and gain-speed tradeoff for self-aligned polysilicon-emitter transistors," *Solid-State Electron.*, vol. 31, 1988.

Zde87 P. J. Zdebel, R. J. Balda, B.-Y. Hwang, V. d. l. Torre, and A. Wagner, "MOSAIC III - A high performance bipolar technology with advanced self-aligned devices", *Proc. IEEE Bipolar Circuits and Technology Meeting*, pp. 172-175, Sept. 1987.

Zie66 A. van der Ziel and D. Agouridis, "The cutoff frequency falloff in UHF transistors at high currents," *Proc. IEEE*, vol. 54, pp. 411-412, Mar. 1966.

BIOGRAPHICAL SKETCH

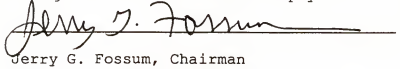
Hanggeun Jeong was born in Chollabukdo, Korea, on March 17, 1955. He received the B.S. degree in electronic engineering from the Seoul National University in 1977 and the M.S. degree in electrical engineering from the Korea Advanced Institute of Science, Seoul, Korea, in 1979. Since January 1983, he has been working toward the Ph.D. degree in electrical engineering at the University of Florida, Gainesville.

From 1979 to 1982 he was with Korea Telecommunications Research Institute, Seoul, Korea, where he was engaged in the development of fiber-optic telecommunication systems.

His current research interests are simulation and design of VLSI semiconductor devices.

He is a member of IEEE.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Jerry G. Fossum, Chairman
Professor of Electrical Engineering

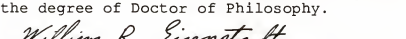
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Fredrik A. Lindholm
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

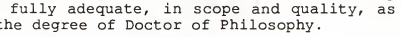
Dorothea E. Burk
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



William R. Eisenstadt
Assistant Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Ulrich H. Kurzweg
Professor of Aerospace Engineering, Mechanics, and Engineering Science

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1989

Hubert A. Benis

Dean, College of Engineering

Dean, Graduate School